Lecture Notes in Mathematics

Marek Biskup
Anton Bovier
Frank den Hollander
Dima Ioffe
Fabio Martinelli
Karel Netočný
Fabio Toninelli

Methods of Contemporary Mathematical Statistical Physics

1970

Editor: Roman Kotecký

Springer

Lecture Notes in Mathematics 1970

Marek Biskup · Anton Bovier · Frank den Hollander · Dima Ioffe · Fabio Martinelli
Karel Netočný · Fabio Toninelli

Methods of Contemporary Mathematical Statistical Physics

Editor:
Roman Kotecký

Springer

Editor
Roman Kotecký
Center for Theoretical Study
Charles University
Jilská 1
100 00, Praha 1
Czech Republic
kotecky@cucc.ruk.cuni.cz

and

Department of Mathematics
University of Warwick
Coventry CV4 7AL
UK

Authors: see List of Contributors

ISBN: 978-3-540-92795-2 e-ISBN: 978-3-540-92796-9
DOI: 10.1007/978-3-540-92796-9

Lecture Notes in Mathematics ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692

Library of Congress Control Number: 2009920693

Mathematics Subject Classification (2000): 82B20, 82B26, 82C20, 60K35, 82C26, 82C35, 82B31, 82B41, 82B44, 82B10, 82C22, 60K05, 82B05, 82C05

Cover picture by Pavel Brázda

Cover design: SPi Publisher Services

Printed on acid-free paper

springer.com

Preface

The Lecture Notes collect seven mini-courses presented at the 5th Prague Summer School on Mathematical Statistical Physics that took place during two weeks of September 2006. As with preceding schools, it was aimed at PhD students and young postdocs. The central theme of the volume is what could be called "mathematics of phase transitions" in diverse contexts. Even though all courses were meant to introduce the reader to recent progress of a particular topic of modern statistical physics, attention has been paid to providing a solid grounding by carefully developing various basic tools.

One of the techniques that led, more than two decades ago, to a series of important outcomes in the theory of phase transitions of lattice models was reflection positivity. Recently it resurfaced and was used to obtain interesting new results in various settings. The lectures of Marek Biskup include a thorough introduction to reflection positivity as well as a review of its recent applications. In addition, it contains a crash course on lattice spin models that is useful as a background for other lectures of the collection.

Also the following two contributions concern equilibrium statistical physics. The lectures of Dmitri Ioffe are devoted to a stochastic geometric reformulation of classical as well as quantum Ising models. A unified approach to the Fortuin-Kasteleyn and random current representations in terms of path integrals is presented.

Statistical mechanics of directed polymers interacting with one-dimensional spatial effects is a topic with various applications in physics and biophysics. The lectures of Fabio Toninelli are devoted to a thorough discussion of the localization/delocalization transition in these models.

Metastability is a topic that has attracted a lot of attention recently. Here it is discussed in the notes of Anton Bovier and Frank den Hollander. The emphasis of the course of Anton Bovier is on a general rigorous framework. It explores how distinct time scales arise in Markov processes and how the metastable exit times can be expressed in terms of capacity, the crucial notion coming from potential theory. The lectures by Frank den Hollander are then devoted to a nontrivial application to metastability in the context of Glauber and Kawasaki dynamics of lattice gases. The main step is the careful evaluation of the relevant capacity in these particular cases.

Readers can have a glimpse of the prolifically developing nonequilibrium realm in the remaining two contributions. The lectures that were presented by Christian Maes and Karel Netočný form a pedagogical account of several recently discussed topics, with an emphasis on general principles.

Facilitated spin models, also known as kinetically constrained spin models, are reflecting important peculiar features of glassy dynamics. The lectures of Fabio Martinelli, submitted here with his coauthors, review mathematical results that contributed to a settlement of questions arising from numerical simulations.

Only one mini-course presented in Prague was not included into the present volume. These are the lectures about computational complexity and phase transitions in combinatorial optimisation presented by Stefan Mertens. The main reason for this ommision is that his presentation was based on the recent monograph written by him and Cris Moore that already covers very well this topic.

The School was organised by Center for Theoretical Study (through the grant MSM 0021620845) with the Institute of Theoretical Computer Science at Charles University providing their beautiful lecture room in the historical centre. It could not have happened without the support of the European Science Foundation under the auspices of the programme Phase Transitions and Fluctuation Phenomena for Random Dynamics in Spatially Extended Systems. But most of all, the success of the School was determined by the lecturers as well as the students who created a pleasant and stimulating atmosphere. We hope that this spirit found its way into the written version of the lecture notes and will be appreciated by the reader.

Prague *Roman Kotecký*
November, 2008

Contents

List of Contributors

Marek Biskup
Department of Mathematics
UCLA, Los Angeles
CA 90095-1555
biskup@math.ucla.edu

Anton Bovier
Institut für Angewandte Mathematik
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstrasse 6
53115 Bonn Germany
bovier@uni-bonn.de

Nicoletta Cancrini
Dip. Matematica, Univ.l'Aquila
1-67100 L'Aquila, Italy
nicoletta.cancrini
@roma1.infn.it

Frank den Hollander
Mathematical Institute
Leiden University
P.O. Box 9512
2300 RA Leiden
The Netherlands
denholla@math.leidenuniv.nl

Dima Ioffe
Faculty of Industrial Engineering and Management, Technion
Haifa, Israel
ieioffe@ie.technion.ac.il

Christian Maes
Instituut voor
Theoretische Fysica
K. U. Leuven, Belgium
Christian.maes@fys.
kuleuven.be

Fabio Martinelli
Dip. Matematica, Univ. Roma Tre, Largo S.L. Murialdo 00146
Roma, Italy
martin@mat.uniroma3.it

Karel Netočný
Institute of Physics AS CR
Prague, Czech Republic
netocny@fzu.cz

Cyril Roberto
Universite Paris-est
L.A.M.A. UMR 8050, 5 bd Descartes, 77454 Marne-la-Vallée
France
cyril.roberto@univ-mlv.fr

Bidzina Shergelashvili
Instituut voor Theoretische Fysica, K. U. Leuven
Belgium

Cristina Toninelli
Laboratoire de Probabilités et Modèles Alèatoires CNRS-UMR 7599 Universités Paris VI-VII 4, Place Jussieu
F-75252 Paris Cedex 05 France
ctoninel@ccr.jussieu.fr

Fabio Toninelli
Laboratoire de Physique
UMR-CNRS 5672, ENS Lyon
46 Allée d'Italie
69364 Lyon Cedex 07, France
fltonine@ens-lyon.fr

Reflection Positivity and Phase Transitions in Lattice Spin Models

Marek Biskup

Department of Mathematics, UCLA, Los Angeles, CA 90095-1555
biskup@math.ucla.edu

Contents

R. Kotecký (ed.), *Methods of Contemporary Mathematical Statistical Physics*, Lecture Notes in Mathematics 1970, DOI: 10.1007/978-3-540-92796-9_1,

1 Introduction

Phase transitions are one of the most fascinating, and also most perplexing, phenomena in equilibrium statistical mechanics. On the physics side, many approximate methods to explain or otherwise justify phase transitions are known but a complete mathematical understanding is available only in a handful of simplest of all cases. One set of tractable systems consists of the so called *lattice spin models*. Originally, these came to existence as simplified versions of (somewhat more realistic) models of crystalline materials in solid state physics but their versatile nature earned them a life of their own in many other disciplines where complex systems are of interest.

The present set of notes describes one successful *mathematical* approach to phase transitions in lattice spin models which is based on the technique of *reflection positivity*. This technique was developed in the late 1970s in the groundbreaking works of F. Dyson, J. Fröhlich, R. Israel, E. Lieb, B. Simon and T. Spencer who used it to establish phase transitions in a host of physically-interesting classical and quantum lattice spin models; most notably, the classical Heisenberg ferromagnet and the quantum XY model and Heisenberg antiferromagnet. Other powerful techniques — e.g., Pirogov-Sinai theory, lace expansion or multiscale analysis in field theory — are available at present that can serve a similar purpose in related contexts, but we will leave their review to experts in those areas.

The most attractive feature of reflection positivity — especially, compared to the alternative techniques — is the simplicity of the resulting proofs. There are generally two types of arguments one can

use: The first one is to derive the so called *infrared bound*, which states in quantitative terms that the fluctuations of the spin variables are dominated by those of a lattice Gaussian free field. In systems with an internal symmetry, this yields a proof of a symmetry-breaking phase transition by way of a spin-condensation argument. Another route goes via the so called *chessboard estimates*, which allow one to implement a Peierls-type argument regardless of whether the model exhibits an internal symmetry or not.

Avid users of the alternative techniques are often quick to point out that the simplicity of proofs has its price: As a rather restrictive condition, reflection positivity applies only to a small (in a well defined sense) class of systems. Fortunately for the technique and mathematical physics in general, the models to which it does apply constitute a large portion of what is interesting for physics, *and* to physicists. Thus, unless one is exclusively after universal statements — i.e., those robust under rather arbitrary perturbations — the route via reflection positivity is often fairly satisfactory.

The spectacular success of reflection positivity from the late 1970s was followed by many interesting developments. For instance, in various joint collaborations, R. Dobrushin, R. Kotecký and S. Shlosman showed how chessboard estimates can be used to prove a phase transition in a class of systems with naturally-defined ordered and disordered components; most prominently, the q-state Potts model for $q \gg 1$. Another neat application came in the papers of M. Aizenman from early 1980s in which he combined the infrared bound with his random-current representation to conclude mean-field critical behavior in the nearest-neighbor Ising ferromagnet above 4 dimensions. Yet another example is the proof, by L. Chayes, R. Kotecký and S. Shlosman, that the Fisher-renormalization scheme in annealed-diluted systems may be substituted by the emergence of an intermediate phase.

These notes discuss also more recent results where their author had a chance to contribute to the field. The common ground for some of these is the use of reflection positivity to provide mathematical justification of "well-known" conclusions from physics folklore. For instance, in papers by N. Crawford, L. Chayes and the present author, the infrared bound was shown to imply that, once a model undergoes a field or energy driven first-order transition in mean-field theory, a similar transition will occur in the lattice model provided the spatial dimension is sufficiently high or the interaction is sufficiently spread-out (but still reflection positive). Another result — due to L. Chayes, S. Starr and the present author — asserts that if a reflection-positive quantum

spin system undergoes a phase transition at intermediate temperatures in its classical limit, a similar transition occurs in the quantum system provided the magnitude of the quantum spin is sufficiently large.

There have also been recent cases where reflection positivity brought a definite end to a controversy that physics arguments were not able to resolve. One instance concerned certain non-linear vector and liquid-crystal models; it was debated whether a transition can occur already in 2 dimensions. This was settled in recent work of A. van Enter and S. Shlosman. Another instance involved spin systems whose (infinite) set of ground states had a much larger set of symmetries than the Hamiltonian of the model; two competing physics reasonings argued for, and against, the survival of these states at low temperatures. Here, in papers of L. Chayes, S. Kivelson, Z. Nussinov and the present author, spin-wave free energy calculations were combined with chessboard estimates to construct a rigorous proof of phase coexistence of only a *finite* number of low-temperature states.

These recent activities show that the full potential of reflection positivity may not yet have been fully exhausted and that the technique will continue to play an important role in mathematical statistical mechanics. It is hoped that the present text will help newcomers to this field learn the essentials of the subject before the need arises to plow through the research papers where the original derivations first appeared.

Organization

This text began as class notes for nine hours of lectures on reflection positivity at the 2006 Prague School and gradually grew into a survey of (part of) this research area. The presentation opens with a review of basic facts about lattice spin models and then proceeds to study two applications of the infrared bound: a spin-condensation argument and a link to mean-field theory. These are followed by the classical derivation of the infrared bound from reflection positivity. The remainder of the notes is spent on applications of a by-product of this derivation, the chessboard estimate, to proofs of phase coexistence.

The emphasis of the notes is on a *pedagogical* introduction to reflection positivity; for this reason we often sacrifice on generality and rather demonstrate the main ideas on the simplest case of interest. To compensate for the inevitable loss of generality, each chapter is endowed with a section "Literature remarks" where we attempt to list the references deemed most relevant to the topic at hand. The notes are closed with a short section on topics that are not covered as well as some open problems that the author finds worthy of some thought.

Acknowledgments

A review naturally draws on the work of many authors. We will do our best to give proper credit to their contribution in the sections "Literature remarks" that appear at the end of each chapter and, of course, in the list of references. I have personally learned a great deal of the subject from Lincoln Chayes with whom I have subsequently coauthored more than half-a-dozen papers some of which will be discussed here. During this time I also benefited from collaborations with N. Crawford, S. Kivelson, R. Kotecký, Z. Nussinov and S. Starr, and from various enlightening discussions involving A. van Enter, B. Nachtergaele and S. Shlosman.

The text would presumably never exist were it not for Roman Kotecký's summer school; I wish to thank him for allowing me to speak on this subject. I am indebted to the participants of the school for comments during the lectures and to T. Bodineau, A. van Enter, E. Lieb and S. Shlosman for suggestions on the first draft of the notes. My presence at the school was made possible thanks to the support from the ESF-program "Phase Transitions and Fluctuation Phenomena for Random Dynamics in Spatially Extended Systems" and from the National Science Foundation under the grant DMS-0505356.

2 Lattice Spin Models: Crash Course

This section prepares the ground for the rest of the course by introducing the main concepts from the theory of Gibbs measures for lattice spin models. The results introduced here are selected entirely for the purpose of these notes; readers wishing a more comprehensive — and in-depth — treatment should consult classic textbooks on the subject.

2.1 Basic Setup

Let us start discussing the setup of the models to which we will direct our attention throughout this course. The basic ingredients are as follows:

- *Lattice*: We will take the d-dimensional hypercubic lattice $\mathbb{Z}^d$ as our underlying graph. This is the graph with vertices at all points in $\mathbb{R}^d$ with integer coordinates and edges between any *nearest neighbor* pair of vertices; i.e., those at Euclidean distance one. We will use $\langle x, y \rangle$ to denote an (unordered) nearest-neighbor pair.

- *Spins*: At each $x \in \mathbb{Z}^d$ we will consider a spin S_x, by which we will mean a random variable taking values in a closed subset Ω of $\mathbb{R}^\nu$, for some $\nu \geq 1$. We will use $S_x \cdot S_y$ to denote a scalar product between S_x and S_y (Euclidean or otherwise).
- *Spin configurations*: For $\Lambda \subset \mathbb{Z}^d$, we will refer to $S_\Lambda := (S_x)_{x\in\Lambda}$ as the spin configuration in Λ. We will be generically interested in describing the statistical properties of such spin configurations with respect to certain (canonical) measures.
- *Boundary conditions*: To describe the law of S_Λ, we will not be able to ignore that some spins are also outside Λ. We will refer to the configuration S_{Λ^c} of these spins as the boundary condition. The latter will usually be fixed and may often even be considered a parameter of the game. When both S_Λ and S_{Λ^c} are known, we will write
$$S := (S_\Lambda, S_{\Lambda^c}) \tag{2.1}$$
to denote their concatenation on all of $\mathbb{Z}^d$.

The above setting incorporates rather varied physical contexts. The spins may be thought of as describing magnetic moments of atoms in a crystal, displacement of atoms from their equilibrium positions or even orientation of grains in nearly-crystalline granular materials.

To define the dynamics of spin systems, we will need to specify the energetics. This is conveniently done by prescribing the *Hamiltonian* which is a function on the spin-configuration space $\Omega^{\mathbb{Z}^d}$ that tells us how much energy each spin configuration has. Of course, to have all quantities well defined we need to fix a *finite* volume $\Lambda \subset \mathbb{Z}^d$ and compute only the energy in Λ. The most general formula we will ever need is
$$H_\Lambda(S) := \sum_{\substack{A\subset\mathbb{Z}^d \text{ finite}\\ A\cap\Lambda\neq\emptyset}} \Phi_A(S) \tag{2.2}$$
where Φ_A is a function that depends only on S_A. To make everything well defined, we require, e.g., that Φ_A is translation invariant and that $\sum_{A\ni 0} \|\Phi_A\|_\infty < \infty$. (The infinity norm may be replaced by some other norm; in particular, should the need arise to talk about unbounded spins.) It is often more convenient to write the above as a formal sum:
$$H(S) := \sum_A \Phi_A(S) \tag{2.3}$$
with the above specific understanding whenever a precise definition is desired.

The energy is not sufficient on its own to define the statistical mechanics of such spin systems; we also need to specify the *a priori measure* on the spins. This will be achieved by prescribing a Borel measure μ_0 on Ω (which may or may not be finite). Before the interaction is "switched on," the spin configurations will be "distributed" according to the product measure, i.e., the *a priori* law of S_Λ is $\bigotimes_{x\in\Lambda} \mu_0(\mathrm{d}S_x)$. The full statistical-mechanical law is then given by a *Gibbs measure* which (in finite volume) takes the general form $\mathrm{e}^{-\beta H(S)} \prod_x \mu_0(\mathrm{d}S_x)$; cf Sect. 2.3 for more details.

2.2 Examples

Here are a few examples of spin systems:

(1) *$O(n)$-model*: Here $\Omega := \mathbb{S}^{n-1} = \{z \in \mathbb{R}^n \colon |z|_2 = 1\}$ with $\mu_0 :=$ surface measure on $\mathbb{S}^{n-1}$. The Hamiltonian is

$$H(S) := -J \sum_{\langle x,y\rangle} S_x \cdot S_y \tag{2.4}$$

where the dot denotes the usual (Euclidean) dot-product in $\mathbb{R}^n$ and $J \geq 0$. (Note that this comes at no loss as the sign of J can be changed by reversing the spins on the odd sublattice of $\mathbb{Z}^d$.)

Note that if $A \in O(n)$ — i.e., A is an n-dimensional orthogonal matrix — then

$$AS_x \cdot AS_y = S_x \cdot S_y \tag{2.5}$$

and so $H(AS) = H(S)$. Since also $\mu_0 \circ A^{-1} = \mu_0$, the model possesses a *global rotation invariance* — with respect to a simultaneous rotation of all spins. (For $n = 1$ this reduces to the invariance under the flip $+1 \leftrightarrow -1$.)

Two instances of this model are known by other names: $n = 2$ is the *rotor model* while $n = 3$ is the (classical) *Heisenberg ferromagnet*.

(2) *Ising model*: Formally, this is the $O(1)$-model. Explicitly, the spin variables σ_x take values in $\Omega := \{-1, +1\}$ with uniform *a priori* measure; the Hamiltonian is

$$H(\sigma) := -J \sum_{\langle x,y\rangle} \sigma_x \sigma_y \tag{2.6}$$

Note that the energy is smaller when the spins at nearest neighbors align and higher when they antialign. (A similar statement holds, of

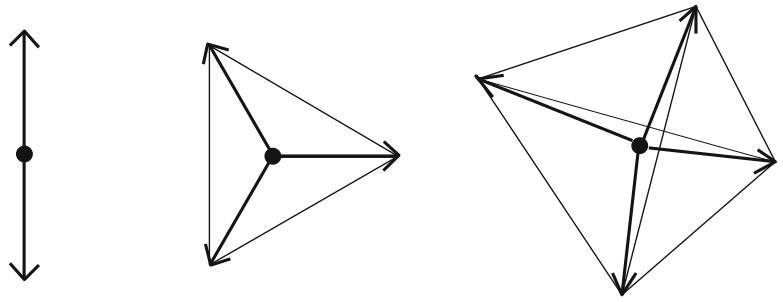
course, for all $O(n)$ models.) This is due to the choice of the sign $J \geq 0$ which makes these models *ferromagnets*.

(3) *Potts model*: This is a generalization of the Ising model beyond two spin states. Explicitly, we fix $q \in \mathbb{N}$ and let σ_x take values in $\{1, \dots, q\}$ (with a uniform *a priori* measure). The Hamiltonian is

$$H(\sigma) := -J \sum_{\langle x,y \rangle} \delta_{\sigma_x, \sigma_y} \tag{2.7}$$

so the energy is $-J$ when σ_x and σ_y "align" and zero otherwise. The $q = 2$ case is the Ising model and $q = 1$ may be related to bond percolation on $\mathbb{Z}^d$ (via the so called *Fortuin-Kasteleyn representation* leading to the so called *random-cluster model*).

It turns out that the Hamiltonian (2.7) can be brought to the form (2.4). Indeed, let Ω denote the set of q points uniformly spread on the unit sphere in $\mathbb{R}^{q-1}$; we may think of these as the vertices of a q-simplex (or a regular q-hedron). The cases $q = 2, 3, 4$ are depicted in this figure:



More explicitly, the elements of Ω are vectors $\hat{\mathrm{v}}_\alpha$, $\alpha = 1, \dots, q$, such that

$$\hat{\mathrm{v}}_\alpha \cdot \hat{\mathrm{v}}_\beta = \begin{cases} 1, & \text{if } \alpha = \beta, \\ -\frac{1}{q-1}, & \text{otherwise.} \end{cases} \tag{2.8}$$

The existence of such vectors can be proved by induction on q. Clearly, if S_x corresponds to σ_x and S_y to σ_y, then

$$S_x \cdot S_y = \frac{q}{q-1} \delta_{\sigma_x, \sigma_y} - \frac{1}{q-1} \tag{2.9}$$

and so the Potts Hamiltonian is to within an additive constant of

$$H(S) := -\tilde{J} \sum_{\langle x,y \rangle} S_x \cdot S_y \tag{2.10}$$

with $\tilde{J} := J\frac{q-1}{q}$. This form, sometimes referred to as *tetrahedral representation*, will be far more useful for our purposes than (2.7).

(4) *Liquid-crystal model*: There are many models that describe certain granular materials known to many of us from digital displays: liquid crystals. A distinguished feature of such materials is the presence of *orientational long-range order* where a majority of the grains align with one another despite the fact that the system as a whole is rotationally invariant. One of the simplest models capturing this phenomenon is as follows: Consider spins $S_x \in \mathbb{S}^{n-1}$ with a uniform *a priori* measure. The Hamiltonian is

$$H(S) := -J \sum_{\langle x,y \rangle} (S_x \cdot S_y)^2 \tag{2.11}$$

The interaction features global rotation invariance and the square takes care of the fact that reflection of any spin does not change the energy (i.e., only the *orientation* rather than the *direction* of the spin matters).

As for the Potts model, the Hamiltonian can again be brought to the form reminiscent of the $O(n)$-model. Indeed, given a spin $S \in \mathbb{S}^{n-1}$ with Cartesian components $S^{(\alpha)}$, $\alpha = 1, \ldots, n$, define an $n \times n$ matrix Q by

$$Q_{\alpha\beta} := S^{(\alpha)} S^{(\beta)} - \frac{1}{n} \delta_{\alpha\beta} \tag{2.12}$$

(The subtraction of the identity is rather arbitrary and more or less unnecessary; its goal is to achieve zero trace and thus reduce the number of independent variables characterizing Q to $n-1$; i.e., as many degrees of freedom as S has.) As is easy to check, if $Q \leftrightarrow S$ and $\tilde{Q} \leftrightarrow \tilde{S}$ are related via the above formula, then

$$\mathrm{Tr}(Q\tilde{Q}) = (S \cdot \tilde{S})^2 - \frac{1}{n} \tag{2.13}$$

Since Q is symmetric, the trace evaluates to

$$\mathrm{Tr}(Q\tilde{Q}) = \sum_{\alpha,\beta} Q_{\alpha\beta} \tilde{Q}_{\alpha\beta} \tag{2.14}$$

which is the canonical scalar product on $n \times n$ matrices. In this language the Hamiltonian takes again the form we saw in the $O(n)$ model.

At this point we pause to remark that all of the above Hamiltonians are of the following rather general form:

$$H(S) = +\frac{1}{2}\sum_{x,y} J_{x,y}|S_x - S_y|^2 \tag{2.15}$$

where (J_{xy}) is a collection of suitable *coupling constants* and $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^n$. This is possible because, in all cases, the (corresponding) norm of S_x is constant and so adding it to the Hamiltonian has no effect on the probability measure. The model thus obtained bears striking similarity to our last example:

(5) *Lattice Gaussian free field*: Let $\Omega := \mathbb{R}$, $\mu_0 :=$ Lebesgue measure and let $\mathsf{P}(x,y)$ be the transition kernel of a symmetric random walk on $\mathbb{Z}^d$; i.e., $\mathsf{P}(x,y) = \mathsf{P}(0,y-x) = \mathsf{P}(0,x-y)$. In this case we will denote the variables by ϕ_x; the Hamiltonian is

$$H(\phi) := \frac{1}{2}\sum_{x,y} \mathsf{P}(x,y)(\phi_y - \phi_x)^2 \tag{2.16}$$

This can be rewritten as

$$H(\phi) = \bigl(\phi, (1-\mathsf{P})\phi\bigr)_{L^2(\mathbb{Z}^d)} =: \mathcal{E}_{1-\mathsf{P}}(\phi,\phi) \tag{2.17}$$

where experts on harmonic analysis of Markov chains will recognize $\mathcal{E}_{1-\mathsf{P}}(\phi,\phi)$ to be the *Dirichlet form* associated with the generator $1-\mathsf{P}$ of the above random walk. In the Gibbs measure, the law of the ϕ_x's will be Gaussian with grad-squared interactions; hence the name of the model.

The sole difference between (2.15) and (2.16) is that, unlike the ϕ_x's, the spins S_x are generally confined to a subset of a Euclidean space and/or their *a priori* measure is not Lebesgue — which will ultimately mean their law is *not* Gaussian. One purpose of this course is to show how this formal similarity can nevertheless be exploited to provide information on the models (2.15).

2.3 Gibbs Formalism

Now we are ready to describe the statistical-mechanical properties of the above models for which we resort to the formalism of Gibbs-Boltzmann distributions. First we define these in finite volume: Given

a finite set $\Lambda \subset \mathbb{Z}^d$ and a boundary condition S_{Λ^c} we define the *Gibbs measure* in Λ to be the probability measure on Ω^Λ given by

$$\mu_{\Lambda,\beta}^{(S_{\Lambda^c})}(\mathrm{d}S_\Lambda) := \frac{\mathrm{e}^{-\beta H_\Lambda(S)}}{Z_{\Lambda,\beta}(S_{\Lambda^c})} \prod_{x\in\Lambda} \mu_0(\mathrm{d}S_x) \tag{2.18}$$

Here $\beta \geq 0$ is the *inverse temperature* — in physics terms, $\beta := \frac{1}{\mathrm{k_B}T}$ where $\mathrm{k_B}$ is the Boltzmann constant and T is the temperature measured in Kelvins — and $Z_{\Lambda,\beta}(S_{\Lambda^c})$ is the normalization constant called the *partition function*.

To extend this concept to infinite volume we have two options:

(1) Consider all possible weak cluster points of the family $\{\mu_{\Lambda,\beta}^{(S_{\Lambda^c})}\}$ as $\Lambda \uparrow \mathbb{Z}^d$ (with the boundary condition possibly varying with Λ) and all convex combinations thereof.
(2) Identify a distinguishing property of Gibbs measures and use it to define infinite volume objects directly.

While approach (1) is ultimately very useful in practical problems, option (2) is more elegant at this level of generality. The requisite "distinguishing property" is as follows:

Lemma 2.1 (DLR condition). *Let $\Lambda \subset \Delta \subset \mathbb{Z}^d$ be finite sets and let $S_{\Delta^c} \in \Omega^{\Delta^c}$. Then (for $\mu_{\Delta,\beta}^{(S_{\Delta^c})}$-a.e. S_{Λ^c}),*

$$\mu_{\Delta,\beta}^{(S_{\Delta^c})}\big(\,\cdot\,\big|S_{\Lambda^c}\big) = \mu_{\Lambda,\beta}^{(S_{\Lambda^c})}(\cdot) \tag{2.19}$$

In simple terms, conditioning the Gibbs measure in Δ on the configuration in $\Delta \setminus \Lambda$, we get the Gibbs measure in Λ with the corresponding boundary condition.

This leads to:

Definition 2.2 (DLR Gibbs measures). *A probability measure on $\Omega^{\mathbb{Z}^d}$ is called an* infinite volume Gibbs measure *for interaction H and inverse temperature β if for all finite $\Lambda \subset \mathbb{Z}^d$ and μ-a.e. S_{Λ^c},*

$$\mu\big(\,\cdot\,\big|S_{\Lambda^c}\big) = \mu_{\Lambda,\beta}^{(S_{\Lambda^c})}(\cdot) \tag{2.20}$$

where $\mu_{\Lambda,\beta}^{(S_{\Lambda^c})}$ is defined using the Hamiltonian H_Λ.

We will use $\mathfrak{G}_\beta$ to denote the set of all infinite volume Gibbs measures at inverse temperature β (assuming the model is clear from the context).

Here are some straightforward, nonetheless important consequences of these definitions:

(1) By Lemma 2.1, any weak cluster point of $(\mu_{\Lambda,\beta}^{(S_{\Lambda^c})})$ belongs to $\mathfrak{G}_\beta$.
(2) By the Backward Martingale Convergence Theorem, if $\Lambda_n \uparrow \mathbb{Z}^d$ and $\mu \in \mathfrak{G}_\beta$, then for μ-a.e. spin configuration S the sequence $\mu_{\Lambda_n,\beta}^{(S_{\Lambda_n^c})}$ has a weak limit, which then belongs to $\mathfrak{G}_\beta$.
(3) $\mathfrak{G}_\beta$ is a convex set (and is closed in the topology of weak convergence). Moreover, $\mu \in \mathfrak{G}_\beta$ is extremal in $\mathfrak{G}_\beta$ iff $\mu_{\Lambda_n,\beta}^{(S_{\Lambda_n^c})} \xrightarrow{\text{w}} \mu$ for μ-almost every spin configuration S.

Similarly direct is the proof of the following "continuity" property:

(4) Let H_n be a sequence of Hamiltonians converging — in the sup-norm on the potentials Φ_A — to Hamiltonian H, and let β_n be a sequence with $\beta_n \to \beta < \infty$. Let μ_n be the sequence of the corresponding Gibbs measures. Then every (weak) cluster point of (μ_n) is an infinite-volume Gibbs measure for the Hamiltonian H and inverse temperature β.

Note that the fact that $\mathfrak{G}_\beta$ is closed and convex ensures that each element can be written as a unique convex combination of extreme points (by the Krein-Millman theorem). The DLR condition permits to extract the corresponding decomposition probabilistically by conditioning on the σ-algebra of tail events.

Now we give a meaning to the terms that are frequently (though sometimes vaguely) employed by physicists:

Definition 2.3 (Phase coexistence). *We say that the model is at* phase coexistence *(or undergoes a* first-order phase transition*) whenever the parameters are such that* $|\mathfrak{G}_\beta| > 1$.

The simplest example where this happens is the Ising model. Let

$$\Lambda_L := \{1, \dots, L\}^d \tag{2.21}$$

and consider the Ising model in Λ_L with all boundary spins set to $+1$. (This is the so called *plus boundary condition*.) As a consequence of stochastic domination — which we will not discuss here — $\mu_{\Lambda_L,\beta}^+$ tends weakly to a measure μ^+ as $L \to \infty$. Similarly, for the *minus* boundary condition, $\mu_{\Lambda_L,\beta}^- \xrightarrow{\text{w}} \mu^-$. It turns out that, in dimensions $d \geq 2$ there exists $\beta_{\text{c}}(d) \in (0, \infty)$ such that

$$\beta > \beta_{\text{c}}(d) \quad \Rightarrow \quad \mu^+ \neq \mu^- \tag{2.22}$$

i.e, the model is at phase coexistence, while for $\beta < \beta_c(d)$, the set of all infinite volume Gibbs measures is a singleton — which means that the model is in the uniqueness regime. One of our goals is to prove similar statements in all of the models introduced above.

2.4 Torus Measures

In the above, we always put a boundary condition in the complement of the finite set Λ. However, it is sometimes convenient to consider other boundary conditions. One possibility is to ignore the existence of Λ^c altogether — this leads to the so called *free boundary condition*. Another possibility is to wrap Λ into a graph without a boundary — typically a torus. This is the case of *periodic* or *torus boundary conditions*.

Consider the torus $\mathbb{T}_L$, which we define as a graph with vertices $(\mathbb{Z}/L\mathbb{Z})^d$, endowed with the corresponding (periodic) nearest-neighbor relation. For nearest-neighbor interactions, the corresponding Hamiltonian is defined easily, but some care is needed for interactions that can be of arbitrary range. If $S \in \Omega^{\mathbb{T}_L}$ we define the *torus Hamiltonian* $H_L(S)$ by

$$H_L(S) := H_{\Lambda_L}(\text{periodic extension of } S \text{ to } \mathbb{Z}^d) \tag{2.23}$$

where we recall $\Lambda_L := \{1, \dots, L\}^d$. For $H(S) := -\frac{1}{2}\sum_{x,y} J_{x,y} S_x \cdot S_y$ we get

$$H_L(S) = -\frac{1}{2} \sum_{x,y \in \mathbb{T}_L} J^{(L)}_{x,y}\, S_x \cdot S_y \tag{2.24}$$

where $J^{(L)}_{x,y}$ are the periodized coupling constants

$$J^{(L)}_{x,y} := \sum_{z \in \mathbb{Z}^d} J_{x,y+Lz} \tag{2.25}$$

The Gibbs measure on $\Omega^{\mathbb{T}_L}$ is then defined accordingly:

$$\mu_{L,\beta}(\mathrm{d}S) := \frac{\mathrm{e}^{-\beta H_L(S)}}{Z_{L,\beta}} \prod_{x \in \mathbb{T}_L} \mu_0(\mathrm{d}S_x) \tag{2.26}$$

where $Z_{L,\beta}$ is the torus partition function. The following holds:

Lemma 2.4. *Every (weak) cluster point of* $(\mu_{L,\beta})_{L \geq 1}$ *lies in* $\mathfrak{G}_\beta$.

Note that there *is* something to prove here because, due to (2.25), the interaction depends on L.

2.5 Some Thermodynamics

Statistical mechanics combines, in its historical development, molecular theory with empirical thermodynamics. Many mathematically rigorous accounts of statistical mechanics thus naturally start the exposition with the notion of the *free energy*. We will need this notion only tangentially — it suffices to think of the free energy as a cumulant generating function — in the proofs of phase coexistence. The relevant statement is as follows:

Theorem 2.5. *For $x \in \mathbb{Z}^d$ let τ_x be the shift-by-x defined by $(\tau_x S)_y := S_{y-x}$. Let $g\colon \Omega^{\mathbb{Z}^d} \to \mathbb{R}$ be a bounded, local function — i.e., one that depends only on a finite number of spins — and recall that $\mu_{L,\beta}$ denote the torus Gibbs measures. Then:*

(1) The limit

$$f(h) := \lim_{L\to\infty} \frac{1}{L^d} \log E_{\mu_{L,\beta}}\biggl\{\exp\Bigl(h \sum_{x\in\mathbb{T}_L} g\circ\tau_x\Bigr)\biggr\} \tag{2.27}$$

exists for all $h \in \mathbb{R}$ and is convex in h.

(2) If $\mu \in \mathfrak{G}_\beta$ is translation invariant, then

$$\frac{\partial f}{\partial h^-}\Big|_{h=0} \le E_\mu(g) \le \frac{\partial f}{\partial h^+}\Big|_{h=0} \tag{2.28}$$

(3) There exist translation-invariant, ergodic measures $\mu^\pm \in \mathfrak{G}_\beta$ such that

$$E_{\mu^\pm}(g) = \frac{\partial f}{\partial h^\pm}\Big|_{h=0} \tag{2.29}$$

Proof of (1), main idea. For compact state-spaces and absolutely-summable interactions, the existence of the limit follows by standard subadditivity arguments. In fact, the limit will exist and will be the same even if we replace $\mu_{L,\beta}$ in (2.27) by any sequence of Gibbs measures in Λ_L with (even L-dependent) boundary conditions. The convexity of f is a consequence of the Hölder inequality applied to the expectation in (2.27). □

Proof of (2). Let $\mu \in \mathfrak{G}_\beta$ be translation invariant and abbreviate

$$Z_L(h) := E_\mu\biggl\{\exp\Bigl(h \sum_{x\in\Lambda_L} g\circ\tau_x\Bigr)\biggr\} \tag{2.30}$$

Since $\log Z_L$ is convex in h (again, by Hölder) we have for any $h > 0$ that

$$\begin{aligned} \log Z_L(h) - \log Z_L(0) &\geq \frac{\partial}{\partial h} \log Z_L(h)\Big|_{h=0} h \\ &= hE_\mu\Big(\sum_{x\in\Lambda_L} g\circ\tau_x\Big) = h|\Lambda_L|E_\mu(g) \end{aligned} \tag{2.31}$$

Dividing by $|\Lambda_L|$, passing to $L \to \infty$ and using that f is independent of the boundary condition, we get

$$f(h) - f(0) \geq hE_\mu(g) \tag{2.32}$$

Divide by h and let $h \downarrow 0$ to get one half of (2.28). The other half is proved analogously. □

Proof of (3). Let $\mathfrak{G}_{\beta,h}$ be the set of Gibbs measures for the Hamiltonian $H - (h/\beta)\sum_x g\circ\tau_x$. A variant of the proof of (2) shows that if $\mu_h \in \mathfrak{G}_{\beta,h}$ is translation-invariant, then

$$\frac{\partial f}{\partial h^-} \leq E_{\mu_h}(g) \leq \frac{\partial f}{\partial h^+} \tag{2.33}$$

In particular, if $h > 0$ we have

$$E_{\mu_h}(g) \geq \frac{\partial f}{\partial h^-} \underset{h>0}{\geq} \frac{\partial f}{\partial h^+}\Big|_{h=0} \tag{2.34}$$

by the monotonicity of derivatives of convex functions. Taking $h \downarrow 0$ and extracting a weak limit from μ_h, we get a Gibbs measure $\mu^+ \in \mathfrak{G}_\beta$ such that

$$E_{\mu^+}(g) \geq \frac{\partial f}{\partial h^+}\Big|_{h=0} \tag{2.35}$$

(The expectations converge because g is a local — and thus continuous, in the product topology — function.) Applying (2) we verify (2.29) for μ^+.

The measure μ^+ is translation invariant and so it remains to show that μ^+ can actually be chosen ergodic. To that end let us first prove that

$$\frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L} g\circ\tau_x \underset{L\to\infty}{\longrightarrow} E_{\mu^+}(g), \quad \text{in } \mu^+\text{-probability} \tag{2.36}$$

The random variables on the left are bounded by the norm of g and have expectation $E_{\mu^+}(g)$ so it suffices to prove that the limsup is no

larger than the expectation. However, if that were not the case, we would have

$$\mu^+\Big(\sum_{x\in\Lambda_L} g\circ\tau_x > \big(E_{\mu^+}(g)+\epsilon\big)|\Lambda_L|\Big) > \epsilon \tag{2.37}$$

for some $\epsilon>0$ and some sequence of L's. But then for all $h>0$,

$$E_{\mu^+}\biggl\{\exp\Bigl(h\sum_{x\in\Lambda_L} g\circ\tau_x\Bigr)\biggr\} \ge \epsilon \mathrm{e}^{|\Lambda_L|h[E_{\mu^+}(g)+\epsilon]} \tag{2.38}$$

In light of the independence of the limit in (1) on the measure we use — as discussed in the sketch of the proof of (1) — this would imply

$$f(h) \ge h\big(E_{\mu^+}(g)+\epsilon\big) \tag{2.39}$$

which cannot hold for all $h>0$ if the right-derivative of f at $h=0$ is to equal $E_{\mu^+}(g)$. Hence (2.36) holds.

By the Pointwise Ergodic Theorem, the convergence in (2.36) actually occurs — and, by (2.36), the limit equals $E_{\mu^+}(g)$ — for μ^+-almost every spin configuration. This implies that the same must be true for any measure in the decomposition of μ^+ into ergodic components. By classic theorems from Gibbs-measure theory, every measure in this decomposition is also in $\mathfrak{G}_\beta$ and so we can choose μ^+ ergodic. □

The above theorem is very useful for the proofs of phase coexistence. Indeed, one can often prove some estimates that via (2.28) imply that f is not differentiable at $h=0$. Then one applies (2.29) to infer the existence of two distinct, ergodic Gibbs measures saturating the bounds in (2.28). Examples of this approach will be discussed throughout these notes.

2.6 Literature Remarks

This section contains only the minimum necessary to understand the rest of the course. For a comprehensive treatment of Gibbs-measure theory, we refer to classic monographs by Ruelle [88], Israel [66], Simon [97] and Georgii [57]. Further general background on statistical mechanics of such systems can be found in Ruelle's "blue" book [89]. The acronym DLR derives from the initials of Dobrushin and the team of Lanford & Ruelle who first introduced the idea of conditional definition of infinite volume Gibbs measures; cf e.g. [32].

The proof of Theorem 2.5 touches upon the subject of *large deviation theory* which provides a mathematical framework for many empirical principles underlying classical thermodynamics. The connection of course appears in various disguises in the textbooks [66, 97, 57]; for expositions dealing more exclusively with large-deviation theory we refer to the books by den Hollander [64], Dembo and Zeitouni [29], and Deuschel and Stroock [30]. For the Pointwise Ergodic Theorem and other facts from *ergodic theory* we refer to the textbooks by, e.g., Krengel [73] and Petersen [86].

Stochastic domination and the FKG inequality — dealing with partial ordering of spin configurations, functions thereof and thus also measures — are discussed in, e.g., Georgii [57] or Grimmett [61]. The proof of (2.22) can alternatively be based on Griffiths' correlation inequalities (Griffiths [60]). The phase coexistence in the Ising model at large β was first proved by Peierls via a contour argument that now bears his name (see Griffiths [59]).

Concerning the historical origin of the various model systems; the $O(n)$ model goes back to Heisenberg (who introduced its quantum version), the Ising model was introduced by Lenz and given to Ising as a thesis problem while the Potts model was introduced by Domb and given to Potts as a thesis problem. Ironically, the $O(1)$-model bears Ising's name even though his conclusions about it were quite wrong! Apparently, Potts was more deserving.

An excellent reference for mathematical physics of liquid crystals is the monograph by de Gennes and Prost [56]; other, more combinatorial models have been considered by Heilmann and Lieb [62] and Abraham and Heilmann [1]. The tetrahedral representation of the Potts model can be found in Wu's review article [106]; the matrix representation of the liquid-crystal model is an observation of Angelescu and Zagrebnov [6]. Gradient fields — of which the GFF is the simplest example — have enjoyed considerable attention in recent years; cf the review articles by Funaki [52], Velenik [104] and Sheffield [92]. Another name for the GFF is *harmonic crystal*.

3 Infrared Bound & Spin-wave Condensation

The goal of this section is to elucidate the significance of the infrared bound — postponing its proof and connection with reflection positivity until Section 5 — and the use thereof in the proofs of symmetry breaking via the mechanism of *spin-wave condensation*. The presence, and absence, of symmetry breaking in the $O(n)$-model with

certain non-negative two-body interactions will be linked to recurrence vs transience of a naturally induced random walk.

3.1 Random Walk Connections

Consider the model with the Hamiltonian

$$H = -\frac{1}{2}\sum_{x,y} J_{xy}\, S_x \cdot S_y \tag{3.1}$$

where the spins S_x are *a priori* independent and distributed according to a measure μ_0 which is supported in a compact set $\Omega \subset \mathbb{R}^\nu$. Assume that the interaction constants satisfy the following requirements:

(I1) $J_{xx} = 0$ and $J_{x,y} = J_{0,y-x}$
(I2) $\sum_x |J_{0,x}| < \infty$ and $\sum_x J_{0,x} = 1$

i.e., the coupling constants are translation invariant, absolutely summable and, for convenience, normalized to have unit strength. We will actually always restrict our attention to the following specific examples:

- *Nearest-neighbor interactions*:

$$J_{x,y} = \begin{cases} \frac{1}{2d}, & \text{if } |x-y| = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

- *Yukawa potentials*:

$$J_{x,y} = C\mathrm{e}^{-\mu|x-y|_1} \tag{3.3}$$

  with $\mu > 0$ and $C > 0$.
- *Power-law decaying potentials*:

$$J_{x,y} = \frac{C}{|x-y|_1^s} \tag{3.4}$$

  with $s > d$ and $C > 0$.

On top of these, we will also permit:

- Any convex combination of the three interactions above (with, of course, positive coefficients).

Note that we are using the ℓ_1-distance (rather than the more natural ℓ_2-distance). This is dictated by our methods of proof (see Lemma 5.5). Also note that the Yukawa potential is in the class of *Kac models* where the coupling constants take the form $J_{x,y} = \gamma^d f(\gamma(x-y))$ for some rapidly decaying function $f\colon \mathbb{R}^d \to [0,\infty)$ with unit L^1-norm.

A unifying feature of all three interactions is that $J_{xy} \geq 0$ which allows us to interpret the coupling constants as the *transition probabilities* of a random walk on $\mathbb{Z}^d$. Explicitly, consider a Markov chain (X_n) on $\mathbb{Z}^d$ with

$$\mathsf{P}_z(X_{n+1} = y | X_n = x) := J_{xy} \tag{3.5}$$

where P_z is the law of the chain started at site z. Of particular interest will be the question whether this random walk is *recurrent* or *transient* — i.e., whether a walk started at the origin returns there infinitely, or only finitely many times. Here is a criterion to this matter:

Lemma 3.1. *Let* $\hat{J}(k) := \sum_x J_{0,x} \mathrm{e}^{\mathrm{i}k\cdot x}$, $k \in [-\pi, \pi]^d$. *Then* (X_n) *is transient if and only if*

$$\int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{(2\pi)^d} \frac{1}{1 - \hat{J}(k)} < \infty \tag{3.6}$$

Proof. Recall that a random walk is transient if and only if the first return time to the origin, $\tau_0 := \inf\{n > 0 \colon X_n = 0\}$, is infinite with a positive probability, i.e., $\mathsf{P}_0(\tau_0 < \infty) < 1$. By the formula $\mathsf{E}_0 N = [1 - \mathsf{P}_0(\tau_0 < \infty)]^{-1}$ — where E_0 is the expectation with respect to P_0 — we thus get that transience is equivalent $\mathsf{E}_0 N < \infty$. To compute the expectation, we note

$$1_{\{X_n = 0\}} = \int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{(2\pi)^d} \mathrm{e}^{\mathrm{i}k\cdot X_n} \tag{3.7}$$

which via $\mathsf{E}_0 \mathrm{e}^{\mathrm{i}k\cdot X_n} = [\mathsf{E}_0 \mathrm{e}^{\mathrm{i}k\cdot X_1}]^n = [\sum_x J_{0,x} \mathrm{e}^{\mathrm{i}k\cdot x}]^n = \hat{J}(k)^n$ implies

$$\mathsf{P}_0(X_n = 0) = \int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{(2\pi)^d} \hat{J}(k)^n \tag{3.8}$$

Summing over $n \geq 0$ yields

$$\mathsf{E}_0 N = \sum_{n \geq 0} \int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{(2\pi)^d} \hat{J}(k)^n = \int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{(2\pi)^d} \frac{1}{1 - \hat{J}(k)} \tag{3.9}$$

whereby the claim follows. (A careful proof of the latter identity requires justification of the exchange of the integral with the infinite sum; one has to represent the LHS as a power series, perform the sum and justify limits via appropriate convergence theorems.) □

As to the above examples, we have:

- *n.n. & Yukawa potentials*: As $k \to 0$,

$$1 - \hat{J}(k) \sim C|k|^2 \tag{3.10}$$

and so (X_n) is transient iff $d \geq 3$.

- *Power-law potentials*: Here as $k \to 0$,

$$1 - \hat{J}(k) \sim C \begin{cases} |k|^{s-d}, & \text{if } s < d+2, \\ |k|^2 \log \frac{1}{|k|}, & \text{if } s = d+2, \\ |k|^2, & \text{if } s > d+2. \end{cases} \tag{3.11}$$

Hence (X_n) is transient iff $d \geq 3$ OR $s < \min\{d+2, 2d\}$.

(Note that the walk with $s < d+2$ has a stable-law tail with index of stability $\alpha = s - d$.) A convex combination of the three coupling constants will lead to a transient walk provided at least one of the interactions involved therein (with non-zero coefficients) is transient.

3.2 Infrared Bound

The principal claim of this section is that the finiteness of the integral in (3.6) is sufficient for the existence of a symmetry-breaking phase transition in many spin systems of the kind (3.1). The reason is the connection of the above random walk to the Gaussian free field (2.16) (GFF) with $\mathsf{P}(x,y) := J_{xy}$. Indeed, consider the field in a square box Λ with, say, zero boundary condition. It turns out that

$$\mathrm{Cov}_\Lambda(\phi_x, \phi_y) = \sum_{n \geq 0} \mathsf{P}_x(X_n = y,\, \tau_{\Lambda^c} = y) =: G_\Lambda(x,y) \tag{3.12}$$

where τ_{Λ^c} is the first exit time of the walk from Λ and G_Λ denotes the so called *Green's function* in Λ. In particular, we have

$$\mathrm{Var}_\Lambda(\phi_0) = G_\Lambda(0,0) \tag{3.13}$$

which, as we will see, tends to the integral (3.6) as $\Lambda \uparrow \mathbb{Z}^d$. Since $E_\Lambda(\phi_0) = 0$ due to our choice of the boundary condition, we conclude

$$\bigl\{\mathrm{Law}(\phi_0)\colon \Lambda \subset \mathbb{Z}^d\bigr\} \text{ is tight iff } (X_n) \text{ is transient} \tag{3.14}$$

Physicists actually prefer to think of this in terms of symmetry breaking: Formally, the Hamiltonian of the GFF is invariant under the

transformation $\phi_x \to \phi_x + c$, i.e., the model possesses a global spin-translation symmetry. The symmetry group is not compact and so, to define the model even in finite volume, the symmetry needs to be broken by boundary conditions. The existence of a limit law for ϕ_0 can be interpreted as the survival of the symmetry breaking in the thermodynamic limit — while non-existence means that the invariance is restored in this limit.

Our goal is to show that qualitatively the same conclusions hold also for the $O(n)$-spin system. Explicitly, we will prove:

Theorem 3.2. *Let (J_{xy}) be one of the 3 interactions above. Then:*

Global rotation symmetry of $O(n)$-model is broken at low temperatures $\iff$ *Random walk driven by (J_{xy}) is transient*

We begin with the proof of the implication $\Longleftarrow$. The principal tool will be our next theorem which, for technical reasons, is formulated for torus boundary conditions:

Theorem 3.3 (Infrared Bound). *Let L be an even integer and consider the model* (3.1) *on torus $\mathbb{T}_L$ with Gibbs measure $\mu_{L,\beta}$. Suppose (J_{xy}) is one of the three interactions above and let*

$$c_{L,\beta}(x) := E_{\mu_{L,\beta}}(S_0 \cdot S_x) \tag{3.15}$$

Define $\hat{c}_{L,\beta}(k) := \sum_{x \in \mathbb{T}_L} c_{L,\beta}(x) \mathrm{e}^{\mathrm{i}k \cdot x}$. Then

$$\hat{c}_{L,\beta}(k) \leq \frac{\nu}{2\beta} \frac{1}{1 - \hat{J}(k)}, \qquad k \in \mathbb{T}_L^\star \setminus \{0\} \tag{3.16}$$

where ν is the dimension of the spin vectors and $\mathbb{T}_L^\star$ is the reciprocal torus, $\mathbb{T}_L^\star := \{\frac{2\pi}{L}(n_1, \dots, n_d) \colon n_i = 0, \dots, L-1\}$.

The proof will require developing the technique of reflection positivity and is therefore postponed to Section 5.

Note that $c_{L,\beta}(x)$ is the *spin-spin correlation* function which, in light of translation invariance of $\mu_{L,\beta}$ is a function of only the spatial displacement of the two spins. The result has the following equivalent formulation: For all $(v_x) \in \mathbb{C}^{\mathbb{T}_L}$ with $\sum_x v_x = 0$,

$$\sum_{x,y \in \mathbb{T}_L} v_x \bar{v}_y E_{\mu_{L,\beta}}(S_0 \cdot S_x) \leq \frac{\nu}{2\beta} \sum_{x,y \in \mathbb{T}_L} v_x \bar{v}_y G_L(x,y) \tag{3.17}$$

where

$$G_L(x,y) := \frac{1}{L^d} \sum_{k \in \mathbb{T}_L^\star \setminus \{0\}} \frac{\mathrm{e}^{\mathrm{i}k\cdot(x-y)}}{1-\hat{J}(k)} \tag{3.18}$$

Observe that the latter is the covariance matrix of the GFF on $\mathbb{T}_L$, projected on the set of configurations with total integral zero (i.e., on the orthogonal complement of constant functions). This is a meaningful object because while the ϕ_x are not really well defined — due to the absence of the boundary — the differences $\phi_y - \phi_x$ are. (These differences are orthogonal to constant functions, of course.) A short formulation of the infrared bound is thus:

The correlation of the spins in models (3.1) *with one of the three interactions above is dominated — as a matrix on the orthogonal complement of constant functions in* $L^2(\mathbb{T}_L)$ *— by the covariance of a GFF.*

This fact is often referred to as *Gaussian domination*.

3.3 Spin-wave Condensation in $O(n)$-model

Having temporarily dispensed with the IRB, we will continue in our original line of thought. Theorem 3.3 implies:

Corollary 3.4 (Spin-wave Condensation). *Suppose* $|S_x| = 1$. *Then*

$$E_{\mu_{L,\beta}}\biggl(\Bigl|\frac{1}{L^d}\sum_{x\in\mathbb{T}_L} S_x\Bigr|^2\biggr) \geq 1 - \frac{\nu}{2\beta} G_L(0,0) \tag{3.19}$$

Proof. Let $\hat{S}_k := \sum_{x\in\mathbb{T}_L} S_x \mathrm{e}^{\mathrm{i}k\cdot x}$ be the Fourier coefficient of the decomposition of (S_x) into the so called *spin waves*. The IRB yields

$$E_{\mu_{L,\beta}}|\hat{S}_k|^2 \leq \frac{\nu}{2\beta}\,\frac{L^d}{1-\hat{J}(k)}, \qquad k \in \mathbb{T}_L^\star \setminus \{0\} \tag{3.20}$$

On the other hand, Parseval's identity along with $|S_x| = 1$ implies

$$\sum_{k\in\mathbb{T}_L^\star} |\hat{S}_k|^2 = L^d \sum_{x\in\mathbb{T}_L} |S_x|^2 = L^{2d} \tag{3.21}$$

The IRB makes no statement about $\hat{S}_0$ so we split it from the rest of the sum:

$$\frac{1}{L^{2d}}|\hat{S}_0|^2 = 1 - \frac{1}{L^{2d}} \sum_{k\in\mathbb{T}_L^\star\setminus\{0\}} |\hat{S}_k|^2 \tag{3.22}$$

Now take expectation and apply (3.20):

$$E_{\mu_{L,\beta}}\Big(\frac{1}{L^{2d}}|\hat{S}_0|^2\Big) \geq 1 - \frac{\nu}{2\beta}\,\frac{1}{L^d}\sum_{k\in\mathbb{T}_L^\star\setminus\{0\}}\frac{1}{1-\hat{J}(k)} \tag{3.23}$$

In light of (3.18), this is (3.19). □

With (3.19) in the hand we can apply the same reasoning as for the GFF: In the transient cases, $G_L(0,0)$ converges to the integral (3.6) and so the right-hand side has a finite limit. By taking β sufficiently large, the limit is actually strictly positive. This in turn implies that the zero mode of the spin-wave decomposition is *macroscopically* populated — very much like the free Bose gas at *Bose-Einstein condensation*. Here is how we pull the corresponding conclusions from $\mathbb{T}_L$ onto $\mathbb{Z}^d$:

Theorem 3.5 (Phase Coexistence in $O(n)$-model). *Consider the $O(n)$-model with $n \geq 1$ and one of the three interactions above. Let*

$$\beta_0 := \frac{n}{2}\int_{[-\pi,\pi]^d}\frac{\mathrm{d}k}{(2\pi)^d}\frac{1}{1-\hat{J}(k)} \tag{3.24}$$

Then for any $\beta > \beta_0$ and any $\theta \in \mathbb{S}^{n-1}$ there exists $\mu_\theta \in \mathfrak{G}_\beta$ which is translation invariant and ergodic such that

$$\frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L} S_x \underset{L\to\infty}{\longrightarrow} m_\star\,\theta, \qquad \mu_\theta\text{-a.s.} \tag{3.25}$$

for some $m_\star = m_\star(\beta) > 0$.

Note that (3.25) implies that the measures μ_θ are mutually singular with respect to one another. Note also that β_0 is finite — and the statement is not vacuous — if and only if the associated random walk is transient.

Proof. Suppose, without loss of generality, that we are in the transient case, i.e., $\beta_0 < \infty$. The idea of the proof is quite simple: We use (3.19) to show that the free energy is not differentiable in an appropriately-chosen external field when this field is set to zero. Then we apply Theorem 2.5 to conclude the existence of the required distinct, ergodic Gibbs measures.

Fix $\theta \in \mathbb{S}^{n-1}$ and define

$$f(h) := \lim_{L\to\infty}\frac{1}{L^d}\log E_{\mu_{L,\beta}}\big(\mathrm{e}^{h\theta\cdot\hat{S}_0}\big) \tag{3.26}$$

(The limit exists by Theorem 2.5.) We want to show that $\frac{\partial}{\partial h^+} f(0) > 0$ (and thus, by symmetry, $\frac{\partial}{\partial h^-} f(0) < 0$). Corollary 3.4 yields

$$E_{\mu_{L,\beta}}\bigl(L^{-2d}|\hat{S}_0|^2\bigr) \geq \frac{\beta-\beta_0}{\beta} + o(1) \tag{3.27}$$

Since $|\hat{S}_0| \leq L^d$, for any $0 < \epsilon < 1$ we have

$$E_{\mu_{L,\beta}}\bigl(L^{-2d}|\hat{S}_0|^2\bigr) \leq \epsilon + \mu_{L,\beta}\bigl(|\hat{S}_0| \geq \epsilon L^d\bigr) \tag{3.28}$$

and so

$$\mu_{L,\beta}\Bigl(|\hat{S}_0| \geq \frac{1}{2}\frac{\beta-\beta_0}{\beta} L^d\Bigr) \geq \frac{1}{2}\frac{\beta-\beta_0}{\beta} + o(1) \tag{3.29}$$

By the $O(n)$ symmetry of the torus measures $\mu_{L,\beta}$, the law of $\hat{S}_0/L^d$ is rotationally invariant with non-degenerate "radius" distribution. This implies

$$\mu_{L,\beta}\Bigl(\theta\cdot\hat{S}_0 \geq \frac{1}{4}\frac{\beta-\beta_0}{\beta} L^d\Bigr) \geq C_n\frac{\beta-\beta_0}{\beta} + o(1) \tag{3.30}$$

where $C_2 := 1/6$ and, in general, $C_n > 0$ is an explicitly obtainable constant. But this means that the exponent in the definition of f is at least $\frac{1}{4}\frac{\beta-\beta_0}{\beta} L^d$ with uniformly positive probability and so

$$\frac{\partial f}{\partial h^+}\Big|_{h=0} \geq \frac{\beta-\beta_0}{4\beta} \tag{3.31}$$

Applying Theorem 2.5, for $\beta > \beta_0$ and any $\theta \in \mathcal{S}^1$ there exists a translation invariant, ergodic Gibbs state $\mu_\theta \in \mathfrak{G}_\beta$ such that

$$E_{\mu_\theta}(\theta\cdot S_x) = \frac{\partial f}{\partial h^+}\Big|_{h=0} > 0 \tag{3.32}$$

Next we need to show that the states μ_θ are actually distinct. The Ergodic Theorem implies

$$\frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L} S_x \underset{L\to\infty}{\longrightarrow} m_\star\tilde{\theta}, \qquad \mu_\theta\text{-a.s.} \tag{3.33}$$

where $\tilde{\theta} \in \mathbb{S}^{n-1}$ and where $m_\star > 0$ is the magnitude of the derivative. Note that, in light of (3.32) and the translation invariance of μ_θ,

$$m_\star\,\theta\cdot\tilde{\theta} = \frac{\partial f}{\partial h^+}\Big|_{h=0} \tag{3.34}$$

The distinctness of μ_θ will follow once we prove (3.25), i.e., $\theta = \tilde{\theta}$. (This is, of course, intuitively obvious because the way we constructed μ_θ indicates that the law of S_x under μ_θ should be biased in the direction of θ.)

Suppose $\tilde{\theta} \neq \theta$. Find a rotation $A \in O(n)$ such that $A\tilde{\theta} = \theta$. Let $\tilde{\mu}$ be the measure such that $E_{\tilde{\mu}}(f(S)) := E_{\mu_\theta}(f(AS))$ for all local functions f. (The existence of such a measure follows from the Kolmogorov Extension Theorem.) Since both the Hamiltonian and the *a priori* measure are $O(n)$-invariant, we have $\tilde{\mu} \in \mathfrak{G}_\beta$. But (3.33) implies $E_{\mu_\theta}(S_x) = m_\star \tilde{\theta}$, and so from (3.34) we have

$$E_{\tilde{\mu}}(\theta \cdot S_x) = E_{\mu_\theta}(\theta \cdot AS_x) = m_\star |\theta|^2 \underset{\tilde{\theta} \neq \theta}{>} m_\star \theta \cdot \tilde{\theta} = \frac{\partial f}{\partial h^+}\Big|_{h=0} \tag{3.35}$$

As $\tilde{\mu}$ is a Gibbs measure, this contradicts the general bounds in Theorem 2.5. Hence, we must have $\theta = \tilde{\theta}$ after all. □

The above statement and proof are formulated for the specific case of the $O(n)$ model. A similar proof will apply the existence of a symmetry-breaking phase transition at low temperatures in the Ising, Potts and the liquid-crystal models in all transient dimensions. As the Ising and Potts model have only a *discrete* set of spin states, a symmetry-breaking transition will occur generally in all dimensions $d \geq 2$. However, this has to be proved by different methods than those employed above (e.g., by invoking chessboard estimates).

Our next goal is to establish the complementary part of Theorem 3.2, i.e., the implication $\Longrightarrow$, which asserts the *absence* of symmetry breaking in the recurrent cases. This argument predates the other direction by 20 years and bears the name of its discoverers:

Theorem 3.6 (Mermin-Wagner Theorem). *Let $n \geq 2$ and consider the $O(n)$-model with non-negative interactions constants $(J_{x,y})$ satisfying the conditions (I1,I2) from Sect. 3.1. Suppose the corresponding random walk is recurrent. Then* every *$\mu \in \mathfrak{G}_\beta$ is invariant under any simultaneous (i.e., homogeneous) rotation of all spins.*

Proof. We will show that the spins can be arbitrarily rotated at an arbitrary small cost of the total energy. (This is why we need $n \geq 2$.) We will have to work with *in*homogeneous rotations to achieve this, so let φ_x be a collection of numbers with $\{x\colon \varphi_x \neq 0\}$ finite and let $\mathrm{i}R$ be a unit element of the Lie algebra $\mathfrak{o}(n)$, i.e., $\mathrm{e}^{\mathrm{i}R\alpha}$ is a rigid rotation of

the unit sphere by angle α about a particular axis. Let ω_φ be the map on configuration space acting on individual spins via

$$\omega_\varphi(S_x) := \mathrm{e}^{\mathrm{i}\varphi_x R} S_x, \qquad x \in \mathbb{Z}^d \tag{3.36}$$

To investigate the effect of such an inhomogeneous rotation on the Hamiltonian, note that

$$\begin{aligned} \omega_\varphi(S_x) \cdot \omega_\varphi(S_y) &= S_x \cdot \mathrm{e}^{\mathrm{i}(\varphi_y - \varphi_x)R} S_y \\ &= S_x \cdot S_y - S_x \cdot [1 - \mathrm{e}^{\mathrm{i}(\varphi_y - \varphi_x)R}] S_y \end{aligned} \tag{3.37}$$

Hence the energy of a configuration in any block $\Lambda \supset \{x\colon \varphi_x \neq 0\}$ transforms as

$$H_\Lambda(\omega_\varphi(S)) = H_\Lambda(S) + \triangle H \tag{3.38}$$

where

$$\triangle H := \frac{1}{2} \sum_{x,y} J_{xy}\, S_x \cdot [1 - \mathrm{e}^{\mathrm{i}(\varphi_y - \varphi_x)R}] S_y \tag{3.39}$$

Using that $\triangle H$ depends only on the portion of the spin configuration in Λ, a simple application of the DLR condition shows that, for any local function f,

$$E_\mu(f \circ \omega_\varphi) = E_\mu(f \mathrm{e}^{-\beta \triangle H}) \tag{3.40}$$

We will now let $\varphi_x \to \alpha$ in a specific way that ensures $\triangle H \to 0$; this will permit us to extract the desired conclusion by limiting arguments.

First we will need to control the φ-dependence of $\triangle H$, so we expand the exponential:

$$\begin{aligned} \triangle H = &- \frac{\mathrm{i}}{2} \sum_{x,y} J_{xy}\, (S_x \cdot RS_y)(\varphi_y - \varphi_x) \\ &+ \frac{1}{4} \sum_{x,y} J_{xy}\, (RS_x \cdot RS_y)(\varphi_y - \varphi_x)^2 + \cdots \end{aligned} \tag{3.41}$$

In the first term we note that the self-adjointness of R — valid by the choice of $\mathrm{i}R$ as an element of the Lie algebra — implies that $J_{xy}\,(S_x \cdot RS_y)$ is symmetric under the exchange of x and y. Since $(\varphi_y - \varphi_x)$ is antisymmetric and finitely supported, the sum is zero. Estimating the remainder by the quadratic term, we thus get

$$|\triangle H| \le C \sum_{x,y} J_{xy} (\varphi_y - \varphi_x)^2 = 2C \mathcal{E}_{1-J}(\varphi, \varphi) \tag{3.42}$$

for some constant $C < \infty$. Here we used that $(RS_x \cdot RS_y)$ is bounded and recalled the definition of the Dirichlet form $\mathcal{E}_{1-J}(\cdot,\cdot)$ of the random walk driven by the $(J_{x,y})$'s.

Our next task will be to control the Dirichlet form under the condition that φ tends to α in every finite set. To that end we fix $0 < R < \infty$ and set

$$\varphi_x := \alpha \, \mathsf{P}_x(\tau_0 < \tau_{\Lambda_R^c}) \tag{3.43}$$

This function equals α at $x = 0$, zero on Λ_R^c and is harmonic (with respect to the generator of the random walk) in $\Lambda_R \setminus \{0\}$. A calculation shows

$$\begin{aligned} \mathcal{E}_{1-J}(\varphi,\varphi) &= \sum_x \varphi_x \sum_y J_{x,y}(\varphi_x - \varphi_y) \\ &\underset{\substack{\text{harmonic or} \\ \text{zero in } \{0\}^c}}{=} \alpha \sum_y J_{0,y}(\alpha - \varphi_y) \end{aligned} \tag{3.44}$$

But the recurrence of the associated random walk implies that $\varphi_y \to \alpha$ as $R \to \infty$ for every y and since the J_{xy}'s are summable, the right-hand side tends to zero by the Dominated Convergence Theorem. Thus $\triangle H \to 0$ as $R \to \infty$ and so applying $R \to \infty$ to (3.40) with the choice (3.43) yields

$$E_\mu(f \circ \omega_\alpha) = E_\mu(f) \tag{3.45}$$

for every continuous local function f. Thereby we conclude that μ is invariant under simultaneous rotation of all spins. □

3.4 Literature Remarks

The content of the entire section is very classical. The Infrared Bound (and its proof based on reflection positivity) was discovered in the seminal work of Fröhlich, Simon and Spencer [50] from 1976 where it was also applied to prove a phase transition in the $O(n)$-model (as well as the isotropic Heisenberg and other models). Dyson, Lieb and Simon [38] showed how to adapt the method to a (somewhat more limited) class of quantum spin models. The technique was further developed and its applications extended in two papers of Fröhlich, Israel, Lieb and Simon [46, 47].

Thanks to the representation (2.12), the proof of a long-range order in the liquid-crystal model, derived by Angelescu and Zagrebnov [6], follows the same route as for the $O(n)$ model. However, the type of long-range order that is concluded for the actual spin system is different. Indeed, let μ be a (weak) cluster point of the torus states. Then

$$\lim_{L\to\infty} \frac{1}{|\Lambda_L|} \sum_{x\in\Lambda_L} \bigl[(S_0 \cdot S_x)^2 - 1/n\bigr] > 0 \tag{3.46}$$

with a positive probability under μ. (The limit exists by the Pointwise Ergodic Theorem.) As μ is $O(n)$-invariant, if S_x were asymptotically independent of S_0 for large x, we would expect $E_\mu(S_0 \cdot S_x)^2 \to 1/n$ as $|x| \to \infty$. Apparently, this is not the case, the direction of S_x remains heavily correlated with the direction of S_0 for arbitrary x, i.e., there is an *orientational* long range order.

Whether or not the $O(n)$ symmetry of the law of S_x is broken is an open (and important) question. (The law of each individual S_x is invariant under the flip $S_x \leftrightarrow -S_x$ and so the magnetization is zero in all states.) As noted before, other models of liquid crystals based on dimers on $\mathbb{Z}^2$ were considered by Heilmann and Lieb [62] and Abraham and Heilmann [1] prior to the work [6]. There an orientational long-range order was proved using chessboard estimates; the question of absence of *complete translational ordering* (i.e., breakdown of translation invariance) remained open.

The Mermin-Wagner theorem goes back to 1966 [82]. Various interesting mathematical treatments and extensions followed [34, 87, 49]; the argument presented here is inspired by the exposition in Simon's book [97]. A fully probabilistic approach to this result, discovered by Dobrushin and Shlosman [34], has the advantage that no regularity conditions need to be posed on the spin-spin interaction provided it takes the form $V(S_x - S_y)$; cf the recent paper by Ioffe, Shlosman and Velenik [65]. Finally, we remark that a beautiful and more in-depth exposition of this material — including quantum systems — was presented at the Prague School in 1996 by Bálint Tóth; his handwritten lecture notes should be available online [103].

The basis of the Mermin-Wagner theorem, as well as its extension, is the *continuum* nature of the spin space. Indeed, in the Ising (and also Potts) model, a low-temperature symmetry breaking occurs even in some recurrent dimensions; e.g., in $d = 2$ for the nearest-neighbor interactions. For what determines the presence and absence of symmetry breaking in $d = 1$, see the work of Aizenman, Chayes, Chayes and Newman [3] and references therein.

The connection with random walk is, of course, made possible by our choice to work with non-negative couplings. However, most of the quantitative conclusions of this section hold without reference to random walks. For detailed expositions of the theory of random walks we recommend the monographs by Spitzer [100] and Lawler [76]; the material naturally appears in most graduate probability textbooks (e.g., Durrett [36]).

It is interesting to note that even in $d = 2$, the nearest-neighbor $O(n)$ model exhibits a phase transition when $n = 2$. Namely, while

the Gibbs state is unique at all $\beta < \infty$, for large β it exhibits power law decay of correlations with β-dependent exponents. This regime is (again, after its discoverers) referred to as the *Kosterlitz-Thouless phase* [70]. A rigorous treatment exits, based on renormalization theory and connection with Coulomb gas, thanks to the pioneering work of Fröhlich and Spencer [51]; see also more recent papers by Dimock and Hurd [31]. This is of much interest in light of recent discovery of new conformally-invariant planar processes — the Schramm-Loewner evolution (a.k.a. SLE). No such phenomenon is expected when $n \geq 3$ though there is a minor opposition to this (e.g., Patrasciou and Seiler [85]).

4 Infrared Bound & Mean-field Theory

In this chapter we will discuss how the infrared bound can be used to control the error in so-called mean-field approximation. Unlike the spin-wave condensation, which is concerned primarily with the infrared — i.e., small-k or large spatial scale — content of the IRB, here will make the predominant use of the finite-k — i.e., short range — part of the IRB. (Notwithstanding, the finiteness of the integral (3.6) is still a prerequisite.)

4.1 Mean-field Theory

Mean-field theory is a versatile approximation technique frequently used by physicists to analyze realistic physical models. We begin by a simple derivation that underscores the strengths, and the shortcomings, of this approach.

Consider a lattice spin model with the usual Hamiltonian (3.1). Pick a translation invariant Gibbs measure $\mu \in \mathfrak{G}_\beta$ and consider the expectation of the spin at the origin. The conditional definition of Gibbs measures (the DLR condition) allows us to compute this expectation by first conditioning on all spins outside the origin. Indeed, the one-spin Gibbs measure is determined by the (one-spin) Hamiltonian

$$H_{\{0\}}(S) = -\sum_x J_{0,x}\, S_0 \cdot S_x = -S_0 \cdot \sum_x J_{0,x} S_x = -S_0 \cdot M_0 \qquad (4.1)$$

where we introduced the shorthand

$$M_0 := \sum_x J_{0,x} S_x \qquad (4.2)$$

Thus, by the DLR,

$$E_\mu(S_0) = E_\mu\biggl(\frac{E_{\mu_0}(S_0\,\mathrm{e}^{\beta S_0\cdot M_0})}{E_{\mu_0}(\mathrm{e}^{\beta S_0\cdot M_0})}\biggr) \tag{4.3}$$

where, abusing the notation slightly, the "inner" expectations are only over S_0 — M_0 acts as a constant here — and the outer expectation is over the spins in $\mathbb{Z}^d \setminus \{0\}$, and thus over M_0.

So far the derivation has been completely rigorous but now comes an *ad hoc* step: We suppose that the random variable M_0 is strongly concentrated about its average so that we can replace it by this average. Denoting

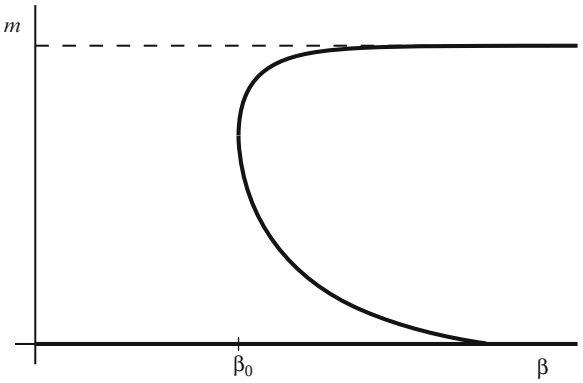
$$m := E_\mu(S_0) \tag{4.4}$$

we thus get that m should be an approximate solution to

$$m = \frac{E_{\mu_0}(S\,\mathrm{e}^{\beta S\cdot m})}{E_{\mu_0}(\mathrm{e}^{\beta S\cdot m})} \tag{4.5}$$

This is the so called *mean-field equation* for the magnetization.

Besides the unjustified step in the derivation, a serious practical problem with (4.5) is that it often has multiple solutions. Indeed, for the set of points (β, m) that obey this equation, one typically gets a picture like this:



Here, for $\beta < \beta_0$, the only solution is $m = 0$ — this is always a solution whenever $E_{\mu_0}(S) = 0$ — but at $\beta = \beta_0$, two new branches appear and coexist over an interval of β's. It is clear that as β varies, the "physical" solution must undergo some sort of jump, but it is not possible to tell where this jump occurs on the basis of equation (4.5) alone. For that one has to go beyond the heuristic derivation presented above.

As is standard, one comes up with an additional "selection" principle that determines which solution is "physical." At the level of classical thermodynamics, this is done by postulating that the solution must minimize an appropriate free energy function. In the choice of this function we will be guided by the fact that there is a proper statistical-mechanical system for which the above derivations can be explicitly validated by way of large-deviation theory. This system is the corresponding model on the *complete graph*.

Consider a graph on N vertices with each pair of vertices joined by an undirected edge. At each vertex $x = 1, \dots, N$ we have a spin S_x with i.i.d. *a priori* law μ_0. Each spin interacts with every other spin; the interaction Hamiltonian is given by

$$H_N(S) := -\frac{1}{2N} \sum_{x,y=1}^{N} S_x \cdot S_y \tag{4.6}$$

The normalization by $1/N$ ensures that the energy grows proportionally to N; the "2" in the denominator compensates for counting each pair of spins twice.

To derive the formula for the free energy function, consider first the cumulant generating function of the measure μ_0,

$$G(h) := \log E_{\mu_0}\big(\mathrm{e}^{h \cdot S}\big), \qquad h \in \mathbb{R}^\nu \tag{4.7}$$

Its Legendre transform,

$$\mathscr{I}(m) := \inf_{h \in \mathbb{R}^\nu} \big[G(h) - h \cdot m\big] \tag{4.8}$$

defines the *entropy* which, according to Cramér's theorem, is the rate of large-deviation decay in

$$\mu_0\bigg(\sum_{x=1}^{N} S_x \approx mN\bigg) = \mathrm{e}^{-N\mathscr{I}(m)+o(N)} \tag{4.9}$$

(The function is infinite outside $\mathrm{Conv}(\Omega)$, the convex hull of Ω and the set of possible values of the magnetization.) Next we inject the energy into the mix and look at the Gibbs measure. To describe what configurations dominate the partition function, and thus the Gibbs measure, we identify the decay rate of the probability

$$\mu_0\Big(\mathrm{e}^{\frac{\beta}{2N}\sum_{x,y=1}^{N} S_x \cdot S_y} 1_{\{\sum_x S_x \approx mN\}}\Big) = \mathrm{e}^{-N\Phi_\beta(m)+o(N)} \tag{4.10}$$

Here the rate function

$$\Phi_\beta(m) := -\frac{\beta}{2}|m|^2 - \mathscr{S}(m) \tag{4.11}$$

is the desired *mean-field free-energy function*. The physical solutions are clearly obtained as the *absolute* minima of $m \mapsto \Phi_\beta(m)$. This is actually completely consistent with (4.5):

Lemma 4.1. *We have*

$$\nabla\Phi_\beta(m) = 0 \qquad \Leftrightarrow \qquad m = \nabla G(\beta m) \tag{4.12}$$

Explicitly, the solutions to (4.5) *are in bijection with the extreme points of* $m \mapsto \Phi_\beta(m)$.

Proof. This is a simple exercise on the Legendre transform. First we note that $\nabla\Phi_\beta(m) = 0$ is equivalent to $\beta m = -\nabla\mathscr{S}(m)$. The convexity of G implies that there is a unique h_m such that $\mathscr{S}(m) = G(h_m) - m \cdot h_m$. Furthermore, h_m depends smoothly on m and we have $\nabla G(h_m) = m$. It is easy to check that then $\nabla\mathscr{S}(m) = -h_m$. Putting this together with our previous observations, we get that

$$\nabla\Phi_\beta(m) = 0 \quad \Leftrightarrow \quad \beta m = h_m \quad \Leftrightarrow \quad m = \nabla G(\beta m) \tag{4.13}$$

It remains to observe that $m = \nabla G(\beta m)$ is a concise way to write (4.5). □

Lemma 4.1 shows that the appearance of multiple solutions to (4.5) coincides with the emergence of secondary local maxima/minima.

4.2 Example: The Potts Model

It is worthwhile to demonstrate the above general formalism on the explicit example of the Potts model. We will work with the tetrahedral representation, i.e., on the spin space $\Omega := \{\hat{\mathrm{v}}_1, \dots, \hat{\mathrm{v}}_q\}$. The mean-field free energy function is best expressed in the parametrization using the *mole fractions*, $x_1, \dots, x_q$, which on the complete graph represent the fractions of all vertices with spins pointing in the directions $\hat{\mathrm{v}}_1, \dots, \hat{\mathrm{v}}_q$, respectively. Clearly,

$$\sum_{i=1}^{q} x_i = 1 \tag{4.14}$$

The corresponding magnetization vector is

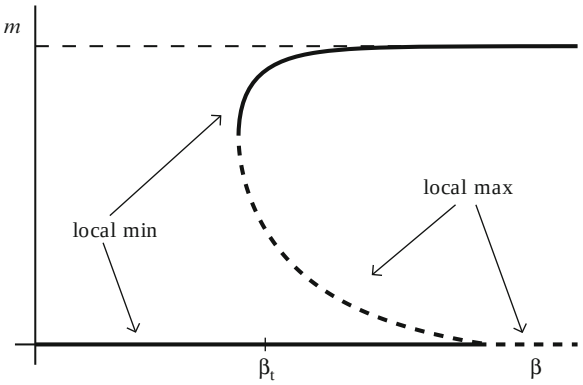
$$m = x_1\hat{\mathrm{v}}_1 + \cdots + x_q\hat{\mathrm{v}}_q. \tag{4.15}$$

In this notation we have

$$\Phi_\beta(m) = \sum_{k=1}^{q} \Big(-\frac{\beta}{2} x_k^2 + x_k \log x_k \Big). \tag{4.16}$$

It is not surprising, but somewhat non-trivial to prove (see [9, Lemma 4.4]) that all interesting behavior of Φ_β occurs "on-axes" that is, the absolute minimizers — and, in fact, all local extrema — of Φ_β occur in the directions of one of the spin states. (Which direction we choose is immaterial as they are related by symmetry.) The following picture shows the qualitative look of the function $\mathfrak{m} \mapsto \Phi_\beta(\mathfrak{m}\hat{\mathrm{v}}_1)$ at four increasing values of β:

Here the function first starts convex and, as β increases, develops a secondary local minimum (plus an inevitable local maximum). For β even larger, the secondary minimum becomes degenerate with the one at $m = 0$ and eventually takes over the role of the global minimum. With these new distinctions, the plot of solutions to the mean-field equation for the magnetization becomes:



Note that the local maximum eventually merges with the local minimum at zero — at which point zero becomes a local maximum. The jump in the position of the global minimum occurs at some $\beta_{\mathfrak{t}}$, which is strictly larger than the point β_0 where the secondary minima/maxima first appear.

4.3 Approximation Theorem & Applications

The goal of this section is to show that, with the help of the IRB, the conclusions of mean-field theory can be given a quantitative form. Throughout we restrict ourselves to interactions of the form (3.1) and the coupling constants being one of the 3 types above.

Definition 4.2. *We say that a measure $\mu \in \mathfrak{G}_\beta$ is a* torus state *if it is either a (weak) cluster point of measures $\mu_{L,\beta}$ or can be obtained from such cluster points by perturbing either β or μ_0 or the inner product between spins.*

The reason for the second half of this definition is that the "operations" thus specified preserve the validity of the IRB. For such states we prove:

Theorem 4.3. *Suppose $|S_x| \le 1$. Let $\mu \in \mathfrak{G}_\beta$ be a translation-invariant, ergodic, torus state and define*

$$m_\star := E_\mu(S_0). \tag{4.17}$$

Let Φ_β be the mean-field free energy function corresponding to this model. Then

$$\Phi_\beta(m_\star) \le \inf_{m \in \mathrm{Conv}(\Omega)} \Phi_\beta(m) + \frac{\nu\beta}{2}\mathcal{I}_d \tag{4.18}$$

where

$$\mathcal{I}_d := \int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{(2\pi)^d} \frac{\hat{J}(k)^2}{1-\hat{J}(k)} \tag{4.19}$$

Note that the integral is finite iff the random walk corresponding to (J_{xy}) is transient. However, unlike for Green's function, $\mathcal{I}_d$ represents the expected number of returns back to the origin after the walk has left the origin. Thus, in strongly transient situations one should expect that $\mathcal{I}_d$ is fairly small. And, indeed, we have the following asymptotics:

- *n.n. interactions*:
$$\mathcal{I}_d \sim \frac{1}{2d}, \qquad d \to \infty. \tag{4.20}$$
- *Yukawa potentials*: If $d \ge 3$,
$$\mathcal{I}_d \le C\mu^d. \tag{4.21}$$

- *Power-law potentials*: If $d \geq 3$ OR $s < \min\{d+2, 2d\}$,

$$\mathcal{I}_d \leq C(s-d). \tag{4.22}$$

Of course, one is able to make the integral small for interactions with power law tails even when s is not too close to d: Just take a mixture of Yukawa and power-law with positive coefficients and let μ be sufficiently small. Within the class of above models, we can rephrase Theorem 4.3 as:

Physical magnetizations nearly minimize
the mean-field free energy function

This is justified because, as it turns out, all relevant magnetizations can be achieved in ergodic torus states. Let us again demonstrate the conclusion on the example of the q-state Potts model:

Theorem 4.4. *Let $q \geq 3$ and suppose that $\mathcal{I}_d \ll 1/q$. Then there is $\beta_{\mathrm{t}} \in (0, \infty)$ and translation-invariant, ergodic measures $\nu_0, \nu_1, \ldots \nu_q \in \mathfrak{G}_{\beta_{\mathrm{t}}}$ such that*

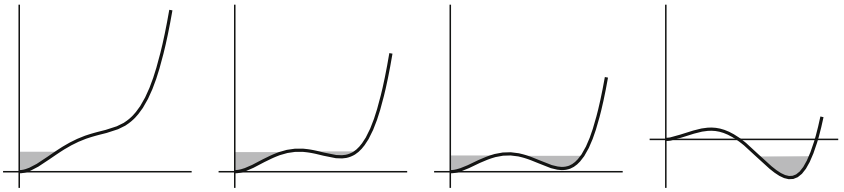
$$|E_{\nu_0}(S_x)| \ll 1 \tag{4.23}$$

and

$$E_{\nu_j}(S_x) = m_\star \, \hat{\mathrm{v}}_j, \qquad j = 1, \ldots, q, \tag{4.24}$$

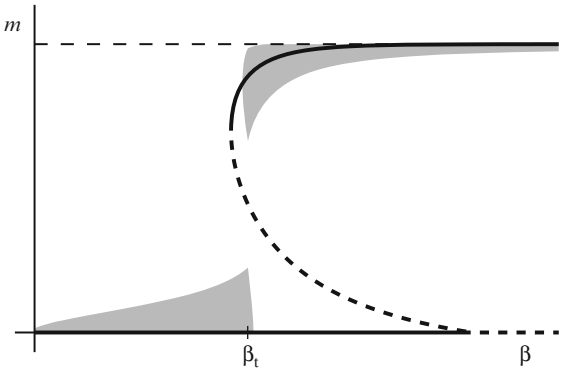
where $m_\star \geq 1/2$. In particular, the 3-state Potts model undergoes a first-order phase transition provided the spatial dimension is sufficiently large.

This result is pretty much the consequence of the pictures in Sect. 4.1. Indeed, including the error bound (4.18), the physical magnetization is confined to the shaded regions:



Thus, once the error is smaller than the "hump" separating the two local minima, there is no way that the physical magnetization can change continuously as the temperature varies. This is seen even more

dramatically once we mark directly into the mean-field magnetization plot the set of values of the magnetization allowed by the inequality (4.18):



(To emphasize the effect, the plots are done for the $q = 10$ state Potts model rather than the most interesting case of $q = 3$.) Notice that the transition is bound to occur rather sharply and very near the mean-field value of β_{t}; explicit error bounds can be derived, but there is no need to state them here.

An additional argument is actually needed to provide a full proof of (4.24). Indeed, we claim that the symmetry breaking happens *exactly* in the direction of one of the spin states while the approximation by mean-field theory only guarantees that the expectation is *near* one of these directions.

Proof of (4.24)*, sketch.* Consider an ergodic Gibbs state μ with $m_\star := E_\mu(S_x) \neq 0$ at inverse temperature β. Given a sample $\sigma = (\sigma_x)$ from μ, at each *unordered* pair $\langle x, y\rangle$ of vertices from $\mathbb{Z}^d$ let

$$\eta_{xy} := 1_{\{\sigma_x = \sigma_y\}} Z_{xy} \tag{4.25}$$

where (Z_{xy}) are *a priori* independent, zero-one valued random variables with

$$\mathbb{P}(Z_{xy} = 1) = 1 - \mathbb{P}(Z_{xy} = 0) := 1 - \mathrm{e}^{-\beta J_{xy}} \tag{4.26}$$

This defines a coupling of μ with a *random cluster measure* — the distribution of the η's — which, by the fact that the extension comes from i.i.d. random variables, is also ergodic.

When $m_\star \neq 0$, the η-marginal features a unique infinite connected component of edges $\langle x, y\rangle$ with $\eta_{xy} = 1$ whose (site) density is proportional to $|m_\star|$. By the construction, the spin variables take a (constant) value on each connected component, which is a.s. unique (by ergodicity) on the infinite one and uniform on the finite ones. Thus, the bias

of the spin distribution comes only from the infinite component and so it points in one of the q spin directions. The claim thus follows. □

4.4 Ideas from the Proofs

A fundamental technical ingredient of the proof is again provided by the IRB, so throughout we will assume one of the three interactions discussed above. However, we will need the following enhanced version:

Lemma 4.5 (IRB Enhanced). *Suppose the random walk driven by the (J_{xy}) is transient and let $G(x,y)$ denote the corresponding Green's function on $\mathbb{Z}^d$. Let $\mu \in \mathfrak{G}_\beta$ be a translation-invariant, ergodic, torus state and let us denote $m_\star := E_\mu(S_0)$. Then for all $(v_x)_{x\in\mathbb{Z}^d} \in \mathbb{C}^{\mathbb{Z}^d}$ with finite support,*

$$\sum_{x,y} v_x \bar{v}_y \, E_\mu\bigl((S_x - m_\star)\cdot(S_y - m_\star)\bigr) \le \frac{\nu}{2\beta} \sum_{x,y} v_x \bar{v}_y G(x,y). \tag{4.27}$$

Proof. The IRB on torus survives weak limits and so we know that, for every (w_x) with finite support *and* $\sum_x w_x = 0$,

$$\sum_{x,y} w_x \bar{w}_y \, E_\mu\bigl(S_x \cdot S_y\bigr) \le \frac{\nu}{2\beta} \sum_{x,y} w_x \bar{w}_y G(x,y) \tag{4.28}$$

where

$$G(x,y) := \lim_{L\to\infty} G_L(x,y) = \int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{(2\pi)^d} \frac{\mathrm{e}^{\mathrm{i}k\cdot(x-y)}}{1-\hat{J}(k)} \tag{4.29}$$

What separates (4.28) from (4.27) are the $m_\star$ terms in the expectation on the left and the absence of the restriction on the sum of v_x. The former is remedied easily; indeed, the restriction $\sum_x w_x = 0$ allows us to put the $m_\star$ terms at no additional cost.

To address the latter issue, suppose (v_x) has finite support but let now $\sum_x v_x$ be arbitrary. Let $\Lambda_L \subset \mathbb{Z}^d$ contain the support of (v_x). To convert to the previous argument, let

$$a_L := \frac{1}{|\Lambda_L|} \sum_x v_x \tag{4.30}$$

and

$$w_x := v_x - a_L 1_{\Lambda_L}(x) \tag{4.31}$$

Note that $\sum_x w_x = 0$. Then

$$\sum_{x,y} w_x \bar{w}_y \, E_\mu\bigl((S_x - m_\star)\cdot(S_y - m_\star)\bigr) = \sum_{x,y} v_x \bar{v}_y \, E_\mu\bigl((S_x - m_\star)\cdot(S_y - m_\star)\bigr) - 2E_\mu\biggl(\Bigl[a_L \sum_{x\in\Lambda_L} (S_x - m_\star)\Bigr] \cdot \Bigl[\sum_y v_y (S_y - m_\star)\Bigr]\biggr) + E_\mu\biggl(\Bigl|a_L \sum_{x\in\Lambda_L} (S_x - m_\star)\Bigr|^2\biggr) \tag{4.32}$$

But ergodicity of μ implies that

$$E_\mu\biggl(\Bigl|\frac{1}{|\Lambda_L|} \sum_{x\in\Lambda_L} (S_x - m_\star)\Bigr|^2\biggr) \underset{L\to\infty}{\longrightarrow} 0 \tag{4.33}$$

and so, by Cauchy-Schwarz, the last two terms in (4.32) converge to zero as $L \to \infty$. Now apply (4.28) and pass to the limit $L \to \infty$ there. A direct calculation (and the Riemann-Lebesgue lemma) shows that

$$\frac{1}{|\Lambda_L|} \sum_{x\in\Lambda_L} G(x,y) \underset{L\to\infty}{\longrightarrow} 0 \tag{4.34}$$

and so the terms involving a_L on the right-hand side of (4.27) suffer a similar fate. This means that the left-hand sides of (4.27–4.28) tend to each other, and same for the right-hand sides. The desired bound (4.27) is thus a limiting version of (4.28). □

Clearly, the restriction to finitely-supported (v_x) is not necessary; instead, one can consider completions of this set in various reasonable norms. The above formulation has an immediate, but rather fundamental, consequence:

Corollary 4.6 (Key Estimate). *Let $\mu \in \mathfrak{G}_\beta$ be an ergodic torus state and let $m_\star := E_\mu(S_x)$. Then we have*

$$E_\mu\biggl(\Bigl|\sum_x J_{0,x} S_x - m_\star\Bigr|^2\biggr) \le \frac{\nu}{2\beta} \mathcal{I}_d. \tag{4.35}$$

Proof. Choose $v_x := J_{0,x}$ and note that with this choice the left-hand side of (4.27) becomes the left-hand side of (4.35). As to the right-hand side of (4.27), we get

$$\frac{\nu}{2\beta} \sum_{x,y} \int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{(2\pi)^d} \frac{\mathrm{e}^{\mathrm{i}k\cdot(x-y)}}{1 - \hat{J}(k)} J_{0,x} J_{0,y} \tag{4.36}$$

Recalling the definition of $\hat{J}(k)$, this yields the desired error term. □

This corollary provides a justification of the *ad hoc* step in the derivation of mean-field theory: Indeed, once $\mathcal{I}_d$ is small, the variance of M_0 is small and so M_0 *is* with high probability close to its average.

The rest of the proof of Theorem 4.3 is based on inequalities linking the mean-field free energy with the actual magnetization of the system; this part of the proof works for general non-negative coupling constants satisfying conditions (I1-I2) from Sect. 3.1. The relevant observations are as follows:

Proposition 4.7. *Let $\mu \in \mathfrak{G}_\beta$ be translation invariant and let $m_\star := E_\mu(S_x)$.*
(1) We have

$$\Phi_\beta(m_\star) \leq \inf_{m\in \mathrm{Conv}(\Omega)} \Phi_\beta(m) + \frac{\beta}{2} \sum_{x\in\mathbb{Z}^d} J_{0,x} \big[E_\mu(S_0 \cdot S_x) - |m_\star|^2 \big] \quad (4.37)$$

(2) Suppose also $J_{0,x} \geq 0$ and $|S_x| \leq 1$. Then

$$\sum_{x\in\mathbb{Z}^d} J_{0,x} \big[E_\mu(S_0 \cdot S_x) - |m_\star|^2 \big] \leq \beta\, E_\mu\Big(\Big| \sum_x J_{0,x}\, S_x - m_\star \Big|^2 \Big) \quad (4.38)$$

Proof of (1). The proof is based on convexity inequalities linking the mean-field free energy and the characteristics of the actual system. Fix $\Lambda \subset \mathbb{Z}^d$ and let Z_Λ be the partition function in Λ. A standard example of such convexity inequality is

$$Z_\Lambda \geq \exp\Big\{ -|\Lambda| \inf_{m\in\mathrm{Conv}(\Omega)} \Phi_\beta(m) + O(\partial\Lambda) \Big\}. \quad (4.39)$$

To prove this we pick m in the (relative) interior of $\mathrm{Conv}(\Omega)$ and define a *tilted* measure

$$\mu_h(\mathrm{d}S) := \mathrm{e}^{h\cdot S - G(h)} \mu_0(\mathrm{d}S) \quad (4.40)$$

with h adjusted so that $E_{\mu_h}(S) = m$. (Such h exists for each m in the relative interior of $\mathrm{Conv}(\Omega)$, by standard arguments for the Legendre transform.) We then get

$$Z_\Lambda = E_{\otimes\mu_h}\big(\mathrm{e}^{-\beta H_\Lambda(S) - h\cdot M_\Lambda + |\Lambda| G(h)} \big) \quad (4.41)$$

where we introduced the shorthand

$$M_\Lambda := \sum_{x\in\Lambda} S_x \quad (4.42)$$

Now apply Jensen to get the expectation into the exponent; the product nature of $\otimes\mu_h$ implies that $E_{\otimes\mu_h}(H_\Lambda(S)) = -|\Lambda|\frac{1}{2}|m|^2 + O(\partial\Lambda)$ and so (4.39) follows by noting that $G(h) - h\cdot m = \mathscr{S}(m)$ due to our choice of h, and subsequently optimizing over all admissible m.

Now fix a general $h \in \mathbb{R}^\nu$ and let μ be a Gibbs measure as specified in the claim. First we note that the DLR condition implies

$$E_\mu(\mathrm{e}^{+\beta H_\Lambda + h\cdot M_\Lambda} Z_\Lambda) = E_{\otimes\mu_0}\bigl(\mathrm{e}^{h\cdot M_\Lambda}\bigr) = \mathrm{e}^{|\Lambda| G(h)} \tag{4.43}$$

The Z_Λ term can be bounded away via (4.39); Jensen's inequality then gives

$$\beta\, E_\mu(H_\Lambda) + |\Lambda|\, h\cdot m_\star - |\Lambda| \inf_{m\in\mathrm{Conv}(\Omega)} \Phi_\beta(m) + O(\partial\Lambda) \le |\Lambda| G(h) \tag{4.44}$$

Next, translation invariance of μ yields

$$E_\mu(H_\Lambda) = -|\Lambda|\frac{1}{2}\sum_x J_{0,x} E_\mu(S_0\cdot S_x) + O(\partial\Lambda) \tag{4.45}$$

and so dividing by Λ and taking $\Lambda \uparrow \mathbb{Z}^d$ along cubes gets us

$$-\frac{\beta}{2}\sum_x J_{0,x} E_\mu(S_0\cdot S_x) \;-\; \inf_{m\in\mathrm{Conv}(\Omega)} \Phi_\beta(m) \le G(h) - h\cdot m_\star \tag{4.46}$$

Optimizing over h turns the right-hand side into $\mathscr{S}(m_\star)$. Adding $\frac{1}{2}|m_\star|^2$ on both sides and invoking (4.11) now proves the claim. □

Proof of (2). Let us return to the notation $M_0 := \sum_x J_{0,x} S_x$. The left-hand side of (4.38) can then be written as $E_\mu(S_0\cdot M_0) - |m_\star|^2$. Since $J_{0,0} = 0$, an application of the DLR condition yields

$$E_\mu(M_0\cdot S_0) = E_\mu\bigl(M_0\cdot\nabla G(\beta M_0)\bigr) \tag{4.47}$$

The DLR condition also implies

$$m_\star = E_\mu(M_0) = E_\mu[\nabla G(\beta M_0)] \tag{4.48}$$

and so we have

$$\begin{aligned} &E_\mu(S_0\cdot M_0) - |m_\star|^2 \\ &\quad = E_\mu\Bigl((M_0 - m_\star)\cdot\bigl(\nabla G(\beta M_0) - \nabla G(\beta m_\star)\bigr)\Bigr) \end{aligned} \tag{4.49}$$

But $|S_x| \le 1$ implies that the Hessian of G is dominated by the identity, $\nabla\nabla G(m) \le \mathrm{id}$ at any $m \in \mathrm{Conv}(\Omega)$ — assuming $J_{x,y} \ge 0$ — and so

$$(M_0 - m_\star) \cdot \big(\nabla G(\beta M_0) - \nabla G(\beta m_\star)\big) \le \beta |M_0 - m_\star|^2 \tag{4.50}$$

by the Mean-Value Theorem. Taking expectations proves (4.38). □

Theorem 4.3 now follows by combining Proposition 4.7 with Corollary 4.6. Interestingly, (4.37) gives

$$\sum_{x \in \mathbb{Z}^d} J_{0,x} E_\mu(S_0 \cdot S_x) \ge |m_\star|^2 \tag{4.51}$$

i.e., the actual energy density always exceeds the mean-field energy density.

4.5 Literature Remarks

The inception of mean-field theory goes back to Curie [28] and Weiss [105]. One of the early connections to the models on the complete graph appears in Ellis' textbook on large-deviation theory [40]. Most of this section is based on the papers of Biskup and Chayes [9] and Biskup, Chayes and Crawford [10]. The Key Estimate had been used before in some specific cases; e.g., for the Ising model in the paper by Bricmont, Kesten, Lebowitz and Schonmann [22] and for the q-state Potts model in the paper by Kesten and Schonmann [69]. Both these works deal with the limit of the magnetization as $d \to \infty$; notwithstanding, no conclusions were extracted for the presence of first-order phase transitions in finite-dimensional systems.

The first-order phase transition in the q-state Potts model has first been proved by Kotecký and Shlosman [69] but the technique works only for extremely large q. The case of small q has been open. The upshot of the present technique is that it replaces q by d or interaction range in its role of a "large parameter." The price to pay is the lack of explicit control over symmetry: We expect that the measure ν_0 in Theorem 4.4 is actually "disordered" and $E_{\nu_0}(S_x) = 0$. This would follow if we knew that the magnetization in the Potts model can be discontinuous only at the percolation threshold — for the Ising model this was recently proved by Bodineau [19] — but this is so far known only in $d = 2$ (or for q very large). The coupling in the proof of (4.24) is due to Edwards and Sokal [39]; for further properties see Grimmett [61] or Biskup, Borgs, Chayes and Kotecky [8]. The uniqueness of the infinite connected component is well known in the nearest-neighbor case from a beautiful argument of Burton and Keane [23]; for the long-range models it has to be supplied by a percolation bound dominating the number of edges connecting a box of side L to its complement.

The requirement $\mathcal{I}_d \ll 1/q$ is actually an embarrassment of the theory as the transition should become more pronounced, and thus easier to control, with increasing q. Thus, even for nearest-neighbor case, we still do not have a dimension in which all $q \geq 3$ state Potts models go first order. (It is expected that this happens already in $d = 3$.) The restriction to transient dimensions is actually not absolutely necessary; cf recent work Chayes [24].

It is natural to ask whether one can say anything about the continuum-q extension of the Potts model, the random cluster model; see Grimmett [61]. Unfortunately, the main condition for proving the IRB, reflection positivity, holds if *and only if* q is integer (Biskup [7]).

Another model for which this method yields a novel result is the liquid-crystal model discussed in Sect. 2.2. Here Angelescu and Zagrebnov [6] proved that symmetry breaking (for the order parameter $\max_\alpha E_\mu[S_x^{(\alpha)}]^2 - 1/n$) occurs at low temperatures by exhibiting spin-wave condensation; cf remarks at the end of Chapter 3. In [9] it has been shown that, for $n \geq 3$, the order parameter undergoes a discontinuous transition at intermediate temperatures; van Enter and Shlosman [41, 42] later proved such transitions in highly non-linear cases. Similar "mean-field driven" first order phase transitions have also been proved for the cubic model [9] and the Blume-Capel model [10].

Once the general theory is in place, the proof of a phase transition for a specific model boils down to the analysis of the mean-field free energy function. While in principle always doable, in practice this may be quite a challenge even in some relatively simple examples. See, e.g., [9, Sect. 4.4] what this requires in the context of the liquid-crystal model.

Finally, we note that the IRB has been connected to mean-field theory before; namely, in the work of Aizenman [2] (cf also Fröhlich [45] and Sokal [98]) in the context of lattice field theories and that of Aizenman and Fernández in the context of Ising systems in either high spatial dimensions [4] or for spread-out interactions [5]. A representative result from these papers is that *the critical exponents in the Ising model take mean-field values above 4 dimensions*. The IRB enters as a tool to derive a one-way bound on the critical exponents. Unfortunately, the full conclusions are restricted to interactions that are reflection positive; a non-trivial extension was obtained recently by Sakai [90] who proved the IRB — and the corresponding conclusions about the critical exponents — directly via a version of the lace expansion.

5 Reflection Positivity

In the last two sections we have made extensive use of the infrared bound. Now is the time to prove it. This will require introducing the technique of reflection positivity which, somewhat undesirably, links long-range correlation properties of the spin models under consideration to the explicit structure of the underlying graph. Apart from the infrared bound, reflection positivity yields also the so called chessboard estimate which we will use extensively in Chapter 6.

5.1 Reflection Positive Measures

We begin by introducing the basic setup for the definition of reflection positivity: Consider the torus $\mathbb{T}_L$ of side L with L even. The torus has a natural reflection symmetry along planes orthogonal to one of the lattice directions. (For that purpose we may think of $\mathbb{T}_L$ as embedded into a continuum torus.) The corresponding "plane of reflection" P has two components, one at the "front" of the torus and the other at the "back." The plane either passes through the sites of $\mathbb{T}_L$ or bisects bonds; we speak of reflections *through sites* or *through bonds*, respectively. The plane splits the torus into two halves, $\mathbb{T}_L^+$ and $\mathbb{T}_L^-$, which are disjoint for reflections through bonds and obey $\mathbb{T}_L^+ \cap \mathbb{T}_L^- = P$ for reflections through sites.

Let $\mathfrak{A}^\pm$ denote the set of all functions $f\colon \Omega^{\mathbb{T}_L} \to \mathbb{R}$ that depend only on the spins in $\mathbb{T}_L^\pm$. Let ϑ denote the reflection operator, $\vartheta\colon \mathfrak{A}^\pm \to \mathfrak{A}^\mp$, which acts on spins via

$$\vartheta(S_x) := S_{\vartheta(x)} \tag{5.1}$$

Clearly, ϑ is a morphism of algebra $\mathfrak{A}^+$ onto $\mathfrak{A}^-$ and $\vartheta^2 = \mathrm{id}$.

Definition 5.1 (Reflection Positivity). *A measure μ on $\Omega^{\mathbb{T}_L}$ is* reflection positive *(RP) with respect to ϑ if*

(1) For all $f, g \in \mathfrak{A}^+$,

$$E_\mu(f\,\vartheta g) = E_\mu(g\,\vartheta f) \tag{5.2}$$

(2) For all $f \in \mathfrak{A}^+$,

$$E_\mu(f\,\vartheta f) \geq 0 \tag{5.3}$$

Note that the above implies that $f, g \mapsto E_\mu(f\,\vartheta g)$ is a positive-semindefinite symmetric bilinear form. Condition (5.2) is usually automatically true — it requires only ϑ-invariance of μ — so it is the second condition that makes this concept non-trivial (hence also the name). Here we first note that the concept is not entirely vacuous:

Lemma 5.2. *The product measure, $\mu = \bigotimes \mu_0$, is RP with respect to all reflections.*

Proof. First consider reflections through bonds. Let $f, g \in \mathfrak{A}^+$. Since $\mathbb{T}_L^+ \cap \mathbb{T}_L^- = \emptyset$, the random variables f and ϑg are independent under μ. Hence,

$$E_\mu(f\,\vartheta g) = E_\mu(f)E_\mu(\vartheta g) = E_\mu(f)E_\mu(g) \tag{5.4}$$

whereby both conditions in Definition 5.1 follow.

For reflections through sites, we note that f and ϑg are independent conditional on S_P. Invoking the reflection symmetry of $\mu(\cdot|S_P)$, we get

$$E_\mu(f\,\vartheta g|S_P) = E_\mu(f|S_P)E_\mu(\vartheta g|S_P) = E_\mu(f|S_P)E_\mu(g|S_P) \tag{5.5}$$

Again the conditions of RP follow by inspection. □

A fundamental consequence of reflection positivity is the Cauchy-Schwarz inequality

$$\bigl[E_\mu(f\,\vartheta g)\bigr]^2 \le E_\mu(f\,\vartheta f)E_\mu(g\,\vartheta g) \tag{5.6}$$

Here is an enhanced, but extremely useful, version of this inequality:

Lemma 5.3. *Let μ be RP with respect to ϑ and let $A, B, C_\alpha, D_\alpha \in \mathfrak{A}^+$. Then*

$$\begin{aligned}&\bigl[E_\mu(\mathrm{e}^{A+\vartheta B+\sum_\alpha C_\alpha\,\vartheta D_\alpha})\bigr]^2\\ &\qquad\le \bigl[E_\mu(\mathrm{e}^{A+\vartheta A+\sum_\alpha C_\alpha\,\vartheta C_\alpha})\bigr]\,\bigl[E_\mu(\mathrm{e}^{B+\vartheta B+\sum_\alpha D_\alpha\,\vartheta D_\alpha})\bigr]\end{aligned} \tag{5.7}$$

Proof. Clearly, in the absence of the $C_\alpha\,\vartheta D_\alpha$ terms, this simply reduces to (5.6). To include these terms we use expansion into Taylor series:

$$\begin{aligned}&E_\mu(\mathrm{e}^{A+\vartheta B+\sum_\alpha C_\alpha\,\vartheta D_\alpha})\\ &\quad=\sum_{n\ge0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n} E_\mu\bigl(\,(\mathrm{e}^A C_{\alpha_1}\dots C_{\alpha_n})\,\vartheta(\mathrm{e}^B D_{\alpha_1}\dots D_{\alpha_n})\bigr)\end{aligned} \tag{5.8}$$

Now we apply (5.6) to the expectation on the right-hand side and then one more time to the resulting sum:

$$
\begin{aligned}
&E_\mu(\mathrm{e}^{A+\vartheta B+\sum_\alpha C_\alpha\,\vartheta D_\alpha}) \\
&\quad\le \sum_{n\ge 0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n}\Big[E_\mu\big((\mathrm{e}^A C_{\alpha_1}\dots C_{\alpha_n})\,\vartheta(\mathrm{e}^A C_{\alpha_1}\dots C_{\alpha_n})\big)^{1/2} \\
&\qquad\qquad \times E_\mu\big((\mathrm{e}^B D_{\alpha_1}\dots D_{\alpha_n})\,\vartheta(\mathrm{e}^B D_{\alpha_1}\dots D_{\alpha_n})\big)^{1/2}\Big] \\
&\quad\le \left(\sum_{n\ge 0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n} E_\mu\big((\mathrm{e}^A C_{\alpha_1}\dots C_{\alpha_n})\,\vartheta(\mathrm{e}^A C_{\alpha_1}\dots C_{\alpha_n})\big)\right)^{1/2} \\
&\qquad \times \left(\sum_{n\ge 0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n} E_\mu\big((\mathrm{e}^B D_{\alpha_1}\dots D_{\alpha_n})\,\vartheta(\mathrm{e}^B D_{\alpha_1}\dots D_{\alpha_n})\big)\right)^{1/2}
\end{aligned}
\tag{5.9}
$$

Resummation via (5.8) now yields the desired expression. □

The argument we just saw yields a fundamental criterion for proving reflection positivity:

Corollary 5.4. *Fix a plane of reflection P and let ϑ be the corresponding reflection operator. Suppose that the torus Hamiltonian takes the form*

$$
-H_L = A + \vartheta A + \sum_\alpha C_\alpha\,\vartheta C_\alpha \tag{5.10}
$$

with $A, C_\alpha \in \mathfrak{A}^+$. Then for all $\beta \ge 0$ the torus Gibbs measure, $\mu_{L,\beta}$, is RP with respect to ϑ.

Proof. The proof is a simple modification of the argument in Lemma 5.3: Fix $f, g \in \mathfrak{A}^+$. Expansion of the exponential term in $\sum_\alpha C_\alpha\,\vartheta C_\alpha$ yields

$$
\begin{aligned}
E_{\mu_{L,\beta}}(f\vartheta g) &= \frac{1}{Z_L}E_{\otimes\mu_0}\big(f(\vartheta g)\,\mathrm{e}^{\beta(A+\vartheta A+\sum_\alpha C_\alpha\,\vartheta C_\alpha)}\big) \\
&= \frac{1}{Z_L}\sum_{n\ge 0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n} E_{\otimes\mu_0}\big((f\mathrm{e}^{\beta A}C_{\alpha_1}\cdots C_{\alpha_n})\,\vartheta(g\mathrm{e}^{\beta A}C_{\alpha_1}\cdots C_{\alpha_n})\big)
\end{aligned}
\tag{5.11}
$$

The conditions of RP for $\mu_{L,\beta}$ are now direct consequences of the fact that the product measure, $\bigotimes\mu_0$, is itself RP (cf Lemma 5.3). □

Now we are ready to check that all 3 interactions that we focused our attention on in previous lectures are of the form in Lemma 5.3 and thus lead to RP torus Gibbs measures:

Lemma 5.5. *For any plane P, the n.n. (ferromagnet) interaction, Yukawa potentials and the power-law decaying potentials, the torus Hamiltonian can be written in the form* (5.10) *for some $A, C_\alpha \in \mathfrak{A}^+$.*

Proof. We focus on reflections through bonds; the case of reflections through sites is analogous. Given P, the terms in the Hamiltonian can naturally be decomposed into three groups: those between the sites in $\mathbb{T}_L^+$, those between the sites in $\mathbb{T}_L^-$ and those involving both halves of the torus:

$$-H_L = \underbrace{\frac{1}{2}\sum_{x,y\in\mathbb{T}_L^+} J_{xy}^{(L)}\, S_x\cdot S_y}_{A} + \underbrace{\frac{1}{2}\sum_{x,y\in\mathbb{T}_L^-} J_{xy}^{(L)}\, S_x\cdot S_y}_{\vartheta A} + \sum_{i=1}^{d} \underbrace{\sum_{\substack{x\in\mathbb{T}_L^+\\ y\in\mathbb{T}_L^-}} J_{xy}^{(L)}\, S_x^{(i)} S_y^{(i)}}_{R_i} \tag{5.12}$$

where we used the reflection symmetry of the $J_{xy}^{(L)}$ to absorb the $1/2$ into the sum at the cost of confining x to $\mathbb{T}_L^+$ and y to $\mathbb{T}_L^-$. The first two terms identify A and ϑA; it remains to show that the R_i-term can be written as $\sum_\alpha C_\alpha\,\vartheta C_\alpha$. We proceed on a case-by-case basis:

Nearest-neighbor interactions: Here

$$R_i = \frac{1}{2d}\sum_{\substack{\langle x,y\rangle\\ x\in\mathbb{T}_L^+\\ y\in\mathbb{T}_L^-}} S_x^{(i)} S_y^{(i)} \tag{5.13}$$

which is of the desired form since $S_y = \vartheta(S_x)$ whenever x and y contribute to the above sum.

Yukawa potentials: We will only prove this in $d = 1$; the higher dimensions are harder but similar. Note that if P passes through the origin and $x \in \mathbb{T}_L^+$ and $y \in \mathbb{T}_L^-$,

$$J_{xy}^{(L)} = C\sum_{n\geq 0} \mathrm{e}^{-\mu(|x|+|y|+nL)} \tag{5.14}$$

Hence,

$$R_i = C\sum_{n\geq 0} \mathrm{e}^{-\mu nL}\Bigl(\sum_{x\in\mathbb{T}_L^+} \mathrm{e}^{-\mu|x|} S_x^{(i)}\Bigr)\Bigl(\sum_{y\in\mathbb{T}_L^-} \mathrm{e}^{-\mu|y|} S_y^{(i)}\Bigr) \tag{5.15}$$

which is of the desired form.

Power-law potentials: Here we note

$$\frac{1}{|x-y|_1^s} = \int_0^\infty \mathrm{d}\mu\, \mu^{s-1}\mathrm{e}^{-\mu|x-y|_1} \tag{5.16}$$

which reduces the problem to the Yukawa case. □

We remark that Corollary 5.3 allows a minor generalization: *if a torus measure μ is RP, and a torus Hamiltonian H_L takes the form* (5.10)*, then also the measure* $\mathrm{e}^{-\beta H_L}\mathrm{d}\mu$ *is RP.* This may seem to be a useful tool for constructing RP measures; unfortunately, we do not know any natural measures other than product measures for which RP can be shown directly.

5.2 Gaussian Domination

Now we are in a position to start proving the infrared bound. First we introduce its integral version known under the name Gaussian domination:

Theorem 5.6 (Gaussian Domination). *Let (J_{xy}) be one of the three interactions above. Fix $\beta \geq 0$ and for $h = (h_x)_{x\in\mathbb{T}_L} \in (\mathbb{R}^\nu)^{\mathbb{T}_L}$ define*

$$Z_L(h) := E_{\bigotimes \mu_0}\biggl(\exp\Bigl\{-\beta \sum_{x,y\in\mathbb{T}_L} J_{xy}^{(L)}\,|S_x - S_y + h_x - h_y|^2\Bigr\}\biggr) \tag{5.17}$$

Then

$$Z_L(h) \leq Z_L(0) \tag{5.18}$$

Proof. Let H_L denote the sum in the exponent. It is easy to check that H_L is of the form

$$-H_L = A + \vartheta B + \sum_\alpha C_\alpha\,\vartheta D_\alpha \tag{5.19}$$

Indeed, for $h \equiv 0$ this is simply Lemma 5.5 as the diagonal terms can always by absorbed into the *a priori* measure. To get $h \not\equiv 0$ we replace S_x by $S_x + h_x$ at each x. This changes the meaning of the original terms A and C_α — and makes them different on the two halves of the torus — but preserves the overall structure of the expression.

A fundamental ingredient is provided by Lemma 5.3 which yields

$$Z_L(h)^2 \leq Z_L(h_+)Z_L(h_-) \tag{5.20}$$

where $h_+ := h$ on $\mathbb{T}_L^+$ and $h_+ := \vartheta h$ on $\mathbb{T}_L^-$, and similarly for h_-. Now let us show how this yields (5.18): Noting that $Z_L(h) \to 0$ whenever any component of h tends to $\pm\infty$, the maximum of $Z_L(h)$ is achieved at some finite h. Let $h^\star$ be a maximizer for which

$$N(h) := \#\bigl\{\langle x, y\rangle\colon h_x \neq h_y\bigr\} \tag{5.21}$$

is the smallest among all maximizers. We claim that $N(h^\star)=0$. Indeed, if $N(h^\star)>0$ then there exists a plane of reflection P through bonds such that P intersects at least one bond $\langle x,y\rangle$ with $h^\star_x \neq h^\star_y$. Observe that then

$$\min\bigl\{N(h^\star_+),\,N(h^\star_-)\bigr\} < N(h^\star) \tag{5.22}$$

Suppose without loss of generality that $N(h^\star_+)<N(h^\star)$. Then the fact that $h^\star$ was a maximizer implies

$$Z_L(h^\star)^2 \le Z_L(h^\star_+)Z_L(h^\star_-) \le Z_L(h^\star_+)Z_L(h^\star) \tag{5.23}$$

which means

$$Z_L(h^\star) \le Z_L(h^\star_+) \tag{5.24}$$

i.e., $h^\star_+$ is also a maximizer. But that contradicts the choice of $h^\star$ by which $N(h^\star)$ was already minimal possible. It follows that $N(h^\star)=0$, i.e., $h^\star$ is a constant. Since $Z(h+c)=Z(h)$ for any constant c, (5.18) follows. □

Now we can finally pay an old debt and prove the infrared bound:

Proof of Theorem 3.3. To ease the notation, we will write throughout

$$\langle \eta,\zeta\rangle := \sum_{x\in\mathbb{T}_L} \eta_x\zeta_x \tag{5.25}$$

to denote the natural inner product on $L^2(\mathbb{T}_L)$. First we note that for any $(\eta_x)\in(\mathbb{R}^\nu)^{\mathbb{T}_L}$,

$$\sum_{x,y\in\mathbb{T}_L} J^{(L)}_{xy}|\eta_x-\eta_y|^2 = \langle \eta, G_L^{-1}\eta\rangle \tag{5.26}$$

where G_L is as in (3.18). (Indeed, in Fourier components, $\hat G_L^{-1}(k) = 1-\hat J(k)$.) As is easy to check,

$$\begin{aligned} Z_L(h) &= E_{\bigotimes\mu_0}\bigl(\mathrm{e}^{-\beta\langle S+h, G_L^{-1}(S+h)\rangle}\bigr) \\ &= Z_L(0)E_{\mu_{L,\beta}}\bigl(\mathrm{e}^{-2\beta\langle h,G_L^{-1}S\rangle-\beta\langle h,G_L^{-1}h\rangle}\bigr) \end{aligned} \tag{5.27}$$

where $\mu_{L,\beta}$ is the torus Gibbs measure. The statement of Gaussian domination (5.18) is thus equivalent to

$$E_{\mu_{L,\beta}}\bigl(\mathrm{e}^{-2\beta\langle h,G_L^{-1}S\rangle}\bigr) \le \mathrm{e}^{\beta\langle h,G_L^{-1}h\rangle} \tag{5.28}$$

We will now use invertibility of G_L to replace $G_L^{-1}h$ by h. This yields

$$E_{\mu_{L,\beta}}\big(\mathrm{e}^{-2\beta\langle h,S\rangle}\big) \le \mathrm{e}^{\beta\langle h,G_L h\rangle} \quad \text{whenever} \quad \sum_{x\in\mathbb{T}_L} h_x = 0 \tag{5.29}$$

where the latter condition comes from the fact that G_L^{-1} annihilates constant functions. Next we expand both sides to quadratic order in h:

$$1 - 2\beta E_{\mu_{L,\beta}}\big(\langle h,S\rangle\big) + \frac{4\beta^2}{2} E_{\mu_{L,\beta}}\big(\langle h,S\rangle^2\big) + \cdots \le 1 + \beta\langle h, G_L h\rangle + \cdots \tag{5.30}$$

Since $E_{\mu_{L,\beta}}(S)$ is constant, $E_{\mu_{L,\beta}}(\langle h,S\rangle) = \langle h, E_{\mu_{L,\beta}}(S)\rangle = 0$ and we thus get

$$E_{\mu_{L,\beta}}\big(\langle h,S\rangle^2\big) \le \frac{1}{2\beta}\,\langle h, G_L h\rangle \tag{5.31}$$

Finally, choose $h_x := v_x \hat{\mathrm{e}}_i$, for some orthonormal basis vectors $\hat{\mathrm{e}}_i$ in $\mathbb{R}^\nu$. This singles out the i-th components of the spins on the left-hand side and has no noticeable effect on the right-hand side (beyond replacing vectors h_x by scalars v_x). Summing the result over $i = 1, \ldots, \nu$ we get the dot product of the spins on the left and an extra factor ν on the right-hand side. □

5.3 Chessboard Estimates

The proof of the infrared bound was based on Lemma 5.3 which boils down to the Cauchy-Schwarz inequality for the inner product

$$f, g \mapsto E_\mu(f\,\vartheta g) \tag{5.32}$$

In this section we will systematize the use of the Cauchy-Schwarz inequality to derive bounds on correlation functions. The key inequality — referred to as the *chessboard estimate* — will turn out to be useful in the proofs of phase coexistence in specific spin systems (even those to which the IRB technology does not apply).

Throughout we will restrict attention to reflections through planes of sites as this is somewhat more useful in applications (except for quantum systems). Pick two integers, $B < L$, such that B divides L and L/B is even. Fixing the origin of the torus, let Λ_B the block corresponding to $\{0, 1, \ldots, B\}^d$ — i.e., the block of side B with lower-left corner at the origin. We may cover $\mathbb{T}_L$ by translates of Λ_B,

$$\mathbb{T}_L = \bigcup_{t\in\mathbb{T}_{L/B}} (\Lambda_B + Bt) \tag{5.33}$$

noting that the neighboring translates share the vertices on the adjacent sides. (This is the specific feature of the setup based on reflections through planes of sites.) The translates are indexed by the sites in a "factor torus" $\mathbb{T}_{L/B}$.

Definition 5.7. *A function $f\colon \Omega^{\mathbb{T}_L} \to \mathbb{R}$ is called a B-block function if it depends only on $\{S_x\colon x \in \Lambda_B\}$. An event $\mathcal{A} \subset \Omega^{\mathbb{T}_L}$ is called a B-block event if $1_{\mathcal{A}}$ is a B-block function.*

Given a B-block function f, and $t \in \mathbb{T}_{L/B}$, we define $\vartheta_t f$ be the reflection of f "into" $\Lambda_B + Bt$. More precisely, for a self-avoiding path on $\mathbb{T}_{L/B}$ connecting Λ_B to $\Lambda_B + Bt$, we may sequentially reflect f along the planes between the successive blocks in the path. The result is a function that depends only on $\{S_x\colon x \in \Lambda_B + Bt\}$. Due to the commutativity of the reflections, this function does not depend on the choice of the path, so we denote it simply by $\vartheta_t f$. Note that since reflections are involutive, $\vartheta^2 = \mathrm{id}$, there are only $2d$ distinct functions one can obtain from f modulo translations.

Theorem 5.8 (Chessboard Estimate). *Suppose μ is RP with respect to all reflections between the neighboring blocks of the form $\Lambda_B + Bt$, $t \in \mathbb{T}_{L/B}$. Then for any B-block functions $f_1, \dots, f_m$, and any* distinct *$t_1, \dots, t_m \in \mathbb{T}_{L/B}$,*

$$E_\mu\Big(\prod_{j=1}^m \vartheta_{t_j} f_j\Big) \le \prod_{j=1}^m \Big[E_\mu\Big(\prod_{t\in\mathbb{T}_{L/B}} \vartheta_t f_j\Big)\Big]^{(B/L)^d} \tag{5.34}$$

Here is a version of this bound for events: If $\mathcal{A}_1, \dots, \mathcal{A}_m$ are B-block events and $t_1, \dots, t_m$ are *distinct* elements of $\mathbb{T}_{L/B}$, then

$$\mu\Big(\bigcap_{j=1}^m \vartheta_{t_j}(\mathcal{A}_j)\Big) \le \prod_{j=1}^m \Big[\mu\Big(\bigcap_{t\in\mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}_j)\Big)\Big]^{(B/L)^d} \tag{5.35}$$

where

$$\vartheta_t(\mathcal{A}) := \{\vartheta_t 1_{\mathcal{A}} = 1\} \tag{5.36}$$

Note that the exponent $(B/L)^d$ is the reciprocal volume of the torus $\mathbb{T}_{L/B}$. (This is consistent with the fact that both expressions transform homogeneously under the scaling $f_j \to \lambda_j f_j$ with $\lambda_j \ge 0$.)

Proof of Theorem 5.8. We will assume throughout that $E_\mu(f\,\vartheta f) = 0$ implies $f = 0$. (Otherwise, one has to factor out the ideal of such

functions and work on the factor space.) We will first address the 1D case; the general dimensions will be handled by induction.

Abbreviate $2n := {}^{L}/_{B}$ and fix a collection of non-zero functions $f_1, \ldots, f_{2n}$. Define a multilinear functional F on the set of B-block functions by

$$F(f_1, \ldots, f_{2n}) := E_\mu\Big(\prod_{t=1}^{2n} \vartheta_t f_t\Big) \tag{5.37}$$

Noting that $F(f_j, \ldots, f_j) > 0$, we also define

$$G(f_1, \ldots, f_{2n}) := \frac{F(f_1, \ldots, f_{2n})}{\prod_{j=1}^{2n} F(f_j, \ldots, f_j)^{\frac{1}{2n}}} \tag{5.38}$$

These objects enjoy a natural cyclic invariance,

$$F(f_1, \ldots, f_{2n}) = F(f_{2n}, f_1, \ldots, f_{2n-1}) \tag{5.39}$$

and, similarly,

$$G(f_1, \ldots, f_{2n}) = G(f_{2n}, f_1, \ldots, f_{2n-1}) \tag{5.40}$$

The definition of G also implies

$$G(f, \ldots, f) = 1 \tag{5.41}$$

Finally, Cauchy-Schwarz along the plane separating f_1 from f_{2n} and f_n from f_{n+1} yields

$$\begin{aligned} G(f_1, \ldots, f_{2n}) \leq\ & G(f_1, \ldots, f_n, f_n, \ldots, f_1)^{1/2} \\ & \times\ G(f_{2n}, \ldots, f_{n+1}, f_{n+1}, \ldots, f_{2n})^{1/2} \end{aligned} \tag{5.42}$$

This will of course be the core estimate of the proof.

The desired claim will be proved if we show that

$$G(f_1, \ldots, f_{2n}) \leq 1 \tag{5.43}$$

i.e., that G is maximized by $2n$-tuples composed of the same function. We will proceed similarly as in the proof of Gaussian Domination: Given a $2n$-tuple of B-block functions, $(f_1, \ldots, f_{2n})$, let $(g_1, \ldots, g_{2n})$ be such that

(1) $g_i \in \{f_1, \ldots, f_{2n}\}$ for each $i = 1, \ldots, 2n$
(2) $G(g_1, \ldots, g_{2n})$ maximizes G over all such choices of $g_1, \ldots, g_{2n}$

(3) $g_1, \dots, g_{2n}$ is minimal in the sense that it contains the longest block (counted periodically) of the form $f_i, f_i, \dots, f_i$, for some $i \in \{1, \dots, 2n\}$.

Let k be the length of this block and, using the cyclic invariance, assume that the block occurs at the beginning of the sequence $g_1, \dots, g_{2n}$, i.e., we have $g_1, \dots, g_k = f_i$ (with $g_{k+1} \neq f_i$ unless $k = 2n$).

We claim that $k = 2n$. Indeed, in the opposite case, $k < 2n$, we must have $g_{2n} \neq f_i$ whereby (5.42) combined with the fact that $(g_1, \dots, g_{2n})$ is a maximizer of G imply

$$\begin{aligned} G(g_1, \dots, g_{2n})^2 &\le G(g_1, \dots, g_n, g_n, \dots, g_1) G(g_{2n}, \dots, g_{n+1}, g_{n+1}, \dots, g_{2n}) \\ &\le G(g_1, \dots, g_n, g_n, \dots, g_1) G(g_1, \dots, g_{2n}) \end{aligned} \tag{5.44}$$

i.e.,

$$G(g_1, \dots, g_{2n}) \le G(g_1, \dots, g_n, g_n, \dots, g_1) \tag{5.45}$$

This means that $(g_1, \dots, g_n, g_n, \dots, g_1)$ is also a legitimate maximizer of G but it has a longer constant block — namely of length at least $\min\{2k, 2n\}$. This is a contradiction and so we must have $k = 2n$ after all. In light of (5.41–5.43), this proves the claim in $d = 1$.

To extend the proof to $d > 1$, suppose that $m = (L/B)^d$ and assume, without loss of generality, that we have one function f_t for each block $\Lambda_B + Bt$. Writing

$$\prod_{t \in \mathbb{T}_{L/B}} \vartheta_t f_t = \prod_{j=1}^{2n} \Biggl(\prod_{\substack{t \in \mathbb{T}_{L/B} \\ t_1 = j}} \vartheta_t f_t \Biggr) \tag{5.46}$$

we may apply the 1D chessboard estimate along the product over j. This homogenizes the product over f_t in the first coordinate direction. Proceeding through all directions we eventually obtain the desired claim. □

The chessboard estimate allows us to bound the probability of simultaneous occurrence of distinctly-placed B-block events in terms of their *disseminated versions* $\bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A})$. The relevant quantities to estimate are thus

$$\mathfrak{z}_L(\mathcal{A}) := \mu\Biggl(\bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}) \Biggr)^{(B/L)^d} \tag{5.47}$$

The set function $\mathcal{A} \mapsto \mathfrak{z}_L(\mathcal{A})$ is not generally additive. However, what matters for applications is that it is subadditive:

Lemma 5.9 (Subadditivity). *Let $\mathcal{A}$ and $\mathcal{A}_1, \mathcal{A}_2, \ldots,$ be a collection of B-block events such that*

$$\mathcal{A} \subset \bigcup_k \mathcal{A}_k \tag{5.48}$$

Then

$$\mathfrak{z}_L(\mathcal{A}) \leq \sum_k \mathfrak{z}_L(\mathcal{A}_k) \tag{5.49}$$

Proof. First we use the subadditivity of μ and (5.48) to get

$$\mathfrak{z}_L(\mathcal{A})^{|\mathbb{T}_{L/B}|} = \mu\Big(\bigcap_{t\in\mathbb{T}_{L/B}} \vartheta_t(\mathcal{A})\Big) \underset{(5.48)}{\leq} \sum_{(k_t)} \mu\Big(\bigcap_{t\in\mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}_{k_t})\Big) \tag{5.50}$$

Next we apply the chessboard estimate

$$\mu\Big(\bigcap_{t\in\mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}_{k_t})\Big) \leq \prod_{t\in\mathbb{T}_{L/B}} \mathfrak{z}_L(\mathcal{A}_{k_t}) \tag{5.51}$$

to each term on the right hand side. Finally we apply the distributive law for sums and products with the result

$$\begin{aligned} \mathfrak{z}_L(\mathcal{A})^{|\mathbb{T}_{L/B}|} &\leq \sum_{(k_t)} \prod_{t\in\mathbb{T}_{L/B}} \mathfrak{z}_L(\mathcal{A}_{k_t}) \\ &= \prod_{t\in\mathbb{T}_{L/B}} \sum_k \mathfrak{z}_L(\mathcal{A}_k) = \Big(\sum_k \mathfrak{z}_L(\mathcal{A}_k)\Big)^{|\mathbb{T}_{L/B}|} \end{aligned} \tag{5.52}$$

Taking the $|\mathbb{T}_{L/B}|$-th root now yields the desired claim. □

Here is how subadditivity $\mathfrak{z}_L$ is generally used in computations: In order to estimate the $\mathfrak{z}_L$-value of an event, we first cover it by the union of a collection of smaller — and, as desired, easier to compute-with — events, then evaluate the $\mathfrak{z}_L$-value for each of them and, finally, add the results.

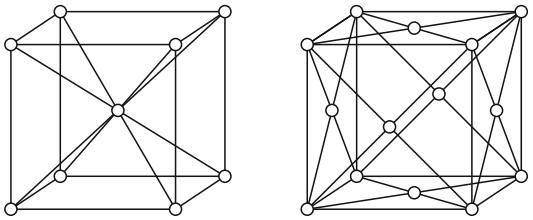
In estimates, we often work with the limiting version,

$$\mathfrak{z}(\mathcal{A}) := \lim_{L\to\infty} \mathfrak{z}_L(\mathcal{A}) \tag{5.53}$$

of this quantity. We may interpret this as a *partition function per site* restricted to event $\mathcal{A}$ on each B-block. The advantage of taking the limit is that it often washes out some annoying finite-size factors and thus provides a more tractable expression to work with. In addition, the limit can be computed using arbitrary — not just periodic — boundary conditions.

5.4 Diagonal Reflections, Other Lattices

The above proof of the chessboard estimate is tailored to the underlying setting of the hypercubic lattice, primarily because of its use of the orthogonality between the principal lattice directions. However, some practical problems may lead us to the consideration of other lattices. Some cases generalize directly, e.g., certain instances of the *body-centered cubic* (BCC) or *face-centered cubic* (FCC) lattices, whose unit cells look respectively as follows:



Both of these are decorations of the cubic lattice in which an extra vertex placed in the center of each unit cube (BCC) or a face (FCC) and is attached by edges to the vertices in its ultimate vicinity.

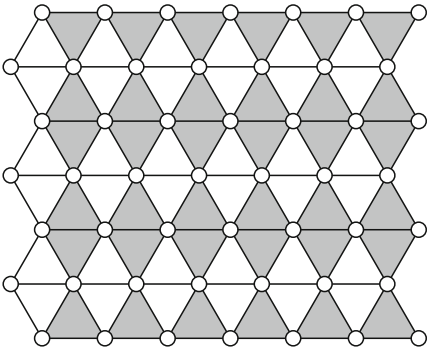
Assuming the interaction (2.15) with J_{xy} non-zero and positive only for adjacent (i.e., nearest-neighbor) pairs of vertices, the torus Gibbs measure is reflection positive for reflections both through and between the planes of sites of $\mathbb{Z}^3$. (A key observation is that the planes between sites of $\mathbb{Z}^3$ contain some of the added vertices but bisect no additional edges.) The strengths of the interactions across the "old" and "new" edges may not even be the same.

In $d = 2$, a corresponding graph is the lattice with a vertex placed in the middle of each square of $\mathbb{Z}^2$ and edges from it to each of the four corners thereof. By the same reasoning, the nearest-neighbor ferromagnetic interaction leads to a reflection positive torus Gibbs measure.

The situation becomes more involved for the *triangular* (two-dimensional) lattice, whose standard embedding into the complex plane $\mathbb{C}$ has vertices

$$m + n\,\mathrm{e}^{\mathrm{i}\pi/3}, \qquad m, n \in \mathbb{Z} \tag{5.54}$$

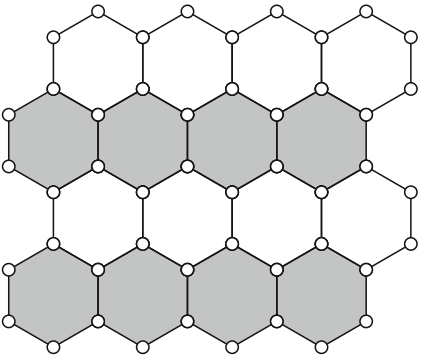
and an edge between any pair of such vertices that differ by a number in the set $\{1, \mathrm{e}^{\mathrm{i}\pi/3}, \mathrm{e}^{\mathrm{i}2\pi/3}\}$. The principal problem with such graphs is how to place a finite piece of this lattice on a torus in a way that gives rise to reflection positive measures. Here is a convenient choice:



with the torus obtained by identifying the vertices on the opposite sides.

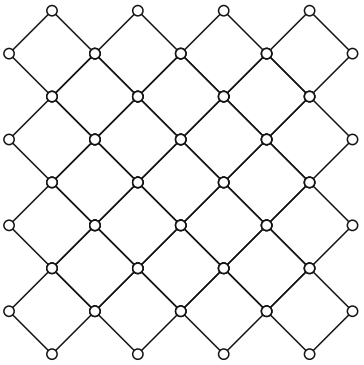
The allowed planes of reflection are all horizontal lines (reflections through sites) and the vertical lines (reflections through both sites and bonds). Again, for ferromagnetic nearest-neighbor interactions, the Gibbs measure with interaction (2.15) is reflection positive. A minor, though annoying, problem occurs in the application of chessboard estimates because the vertical lines of reflections actually cut through triangles. A solution is to focus only on those events that lie either on white or on gray triangles in the above picture and use reflection only with respect to vertical lines that do not cut through the chosen triangles.

A completely analogous situation occurs for the *honeycomb* lattice. Here we consider the domain of the form



and wrap it into a torus by identifying the vertices on the opposite side. Again, for nearest-neighbor ferromagnetic interactions, the resulting Gibbs measure is reflection positive with respect to reflections in vertical lines on sites and horizontal lines between sites. In the application of chessboard estimates to a collection of "hexagon events," we only use every other horizontal and vertical reflections to a corresponding subset of these events; e.g., those sitting on the shaded hexagons.

A final case of interest is that of *diagonal* reflections in $\mathbb{Z}^d$. In $d = 2$, this is achieved by wrapping the domain of the form



periodically into a torus. Reflections in the horizontal and vertical lines of sites — the diagonals — are now symmetries of this graph; for nearest-neighbor interactions (of any sign) the corresponding torus Gibbs measure is reflection positive.

The advantage of the diagonal torus is that it permits the use of reflection positivity on collections of "bond events," i.e., those associated with pairs of nearest-neighbor spins. Subsequent applications of chessboard estimates disseminate a single-bond event over the entire torus. This, in turn, helps in estimates of the quantity $\mathfrak{z}(\mathcal{A})$ whenever $\mathcal{A}$ is an event depending on a single square that is itself an intersection of bond events:

Lemma 5.10. *Given a unit cube in $\mathbb{Z}^d$, let $\mathcal{A}_b$, with b running over all of the $c_d := d2^{d-1}$ edges in this cube, be a collection of bond events. Then*

$$\mathfrak{z}\Biggl(\bigcap_b \mathcal{A}_b\Biggr) \le \prod_b \mathfrak{z}(\mathcal{A}_b)^{1/c_d} \tag{5.55}$$

Here $\mathfrak{z}(\mathcal{A}_b)$ is the partition function per site restricted to configurations such that $\mathcal{A}_b$, or its corresponding reflection, occurs at all edges of $\mathbb{Z}^d$.

Proof. Let us first focus on $d = 2$. The key fact is that the partition function per site, $\mathfrak{z}(\mathcal{A})$, does not depend on what boundary conditions were used to define it. So, in order to compute $\mathfrak{z}$ of the intersection event, we may first wrap the square lattice into the diagonal torus, and disseminate the bond events before passing to the $L \to \infty$. As there are $c_2 = 4$ edges in each lattice square, there is an extra power of $1/4$.

In $d > 2$, we perform the same by singling out two lattice directions and wrapping the torus diagonally in these, and regularly in the

remaining ones. This homogenizes the event in two lattice directions. Proceeding by induction, the claim follows. □

5.5 Literature Remarks

The material of this section is entirely classical; a possible exception is Lemma 5.9 which seems to have been formulated in the present form only relatively recently [13]. The use of reflection positivity goes back to the days of constructive quantum field theory (namely, the Osterwalder-Schrader axioms [84]) where RP was a tool to obtain a sufficiently invariant — and natural — inner product. The use in statistical mechanics was initiated by the work of Fröhlich, Simon and Spencer [50] (infrared bound) and Fröhlich and Lieb [48] (chessboard estimates). The theory was further developed in two papers by Fröhlich, Israel, Lieb and Simon [46, 47]. There have been a couple of nice reviews of this material, e.g., by Shlosman [95] and in Georgii [57].

All use of reflection positivity in these notes is restricted to one of the three interactions introduced in Chapter 3. Various generalizations beyond these are possible. For instance, the n.n. interaction of strength J may be accompanied by a n.n.n. interaction of strength λ — including negative values — and the result is still RP provided $J \geq 2(d-1)|\lambda|$. For reflections through planes of sites, we may even allow any sort of interactions involving the spins in a given lattice cube. (This exhausts all finite range interactions; any longer range RP interactions are automatically infinite range.) Many other examples are discussed, e.g., in [46, page 32].

Notwithstanding our decision to restrict attention only to three specific interactions, the set of reflection positive interactions is not so small as it may appear. Indeed, in the class of translation and rotation invariant coupling constants, letting

$$F(x_1, \dots, x_d) := J_{0,x} \tag{5.56}$$

we check that a sufficient conditions for RP is that the matrix

$$(x, y) \mapsto F(x_1 + y_1, x_2 - y_2, \dots, x_d - y_d) 1_{\{x_1 > 0\}} 1_{\{y_1 > 0\}} \tag{5.57}$$

is positive semidefinite. (See (5.15) for a specific case of this.) By Shur's Theorem — namely that if (a_{ij}) and (b_{ij}) are positive semidefinite matrices, then so is $(a_{ij}b_{ij})$ — we thus know that if $J^{(1)}$ and $J^{(2)}$ are two collections of RP couplings, then also the collection $J^{(1)}_{xy} J^{(2)}_{xy}$ is RP. In particular, the set of RP couplings is closed under taking products.

The situation on other lattices is discussed in [47]; the use of diagonal reflections goes back to [95]. We caution the reader that it is rather easy to make a mistake in this context. For instance, the regularly-wrapped $(L \times L)$ torus in $\mathbb{Z}^2$ is also symmetric with respect to all of the diagonal reflections. However, for diagonal reflection on direct torus it is not possible to define *two* components of the "plane of reflection" so that the reflection in one leaves the other intact. So we cannot simultaneously use both direct and diagonal reflections, and this prevents a direct proof of (5.55) in finite volume. (This error appeared in [17, eq. 4.39] though, as shown in Lemma 5.10, all differences get washed out in the thermodynamic limit.)

Gaussian domination appears in a rather different context as the celebrated *Brascamp-Lieb inequality*. Consider the measure on $\mathbb{R}^n$ of the form

$$\mu(\mathrm{d}x) := Z^{-1}\mathrm{e}^{-V(x)}\mathrm{d}x \tag{5.58}$$

with V smooth and strictly convex. Let $V''(x)$ be the Hessian, i.e., an $n \times n$ matrix of all second derivatives of V. Then for each smooth f with compact support,

$$E_\mu(f^2) - (E_\mu f)^2 \le \int \mathrm{d}x \, \bigl\langle (V'')^{-1}\nabla f(x), \nabla f(x)\bigr\rangle \tag{5.59}$$

where $\langle\cdot,\cdot\rangle$ denote the n-dimensional Euclidean inner product. In particular, if Q is a positive definite $n \times n$ matrix that dominates the Hessian from below at all x, then the correlations of μ are dominated by those of the Gaussian measure with covariance $2Q^{-1}$. This is, unfortunately, not very useful in the analysis of the Gibbs measures for general lattice spin systems as these are generally not of the required form — e.g., because the restriction to a specific spin-space (a unit sphere for the Heisenberg model) cannot be approximated by convex functions.

6 Applications of Chessboard Estimates

In this section we will apply the technique of chessboard estimates to obtain proofs of phase coexistence in some lattice spin models. The arguments will be carried out in detail only for one rather simple example. For more sophisticated systems we present only the important ideas. Details, anyway, can be found in the corresponding papers.

6.1 Gaussian Double-well Model

Here we will demonstrate the use of chessboard estimates on the model of a Gaussian free-field model in a non-quadratic, double-well on-site potential. The Hamiltonian takes the general form

$$\beta H(\phi) := \beta \sum_{\langle x,y \rangle} (\phi_x - \phi_y)^2 + \sum_x V(\phi_x) \tag{6.1}$$

where $\phi_x \in \mathbb{R}$ with *a priori* measure given by the Lebesgue measure, and V is a potential. Note that β has been incorporated into the Hamiltonian in such a way that the on-site potential remains independent of it.

The most well known example of such systems is $V(\phi) := \frac{\kappa}{2}\phi^2$ with $\kappa > 0$ which is known as the *massive* Gaussian free field. This case can of course be treated completely explicitly; e.g., on the torus the corresponding Gaussian measure on (ϕ_x) is zero-mean with covariance

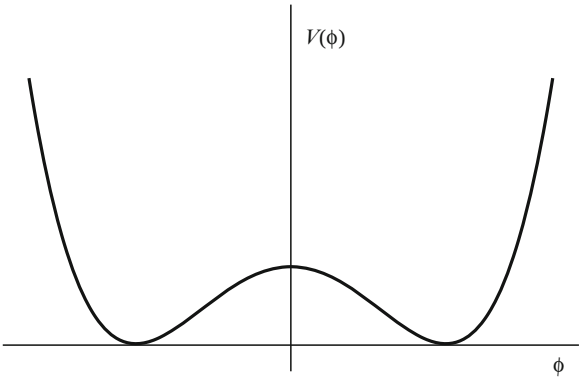
$$\mathrm{Cov}(\phi_x, \phi_y) = \sum_{k \in \mathbb{T}_L^\star} \frac{\mathrm{e}^{\mathrm{i}k\cdot(x-y)}}{\beta \widehat{D}(k) + \kappa} \tag{6.2}$$

where $\widehat{D}(k)$ is the Fourier transform of the torus (discrete) Laplacian,

$$\widehat{D}(k) := \sum_{j=1}^{d} |1 - \mathrm{e}^{\mathrm{i}k_j}|^2 \tag{6.3}$$

Note that the inclusion of the mass, $\kappa > 0$ — more precisely, κ is the mass squared — makes the covariance regular even for the zero mode $k = 0$.

We will look at a modification of this case when V takes the form



In fact, we will be even more specific and assume that V is simply given by

$$\mathrm{e}^{-V(\phi)} := \mathrm{e}^{-\frac{\kappa}{2}(\phi-1)^2} + \mathrm{e}^{-\frac{\kappa}{2}(\phi+1)^2} \tag{6.4}$$

It is easy to check that, for κ sufficiently large, V defined using this formula looks as in the figure. The reason for assuming (6.4) is the possibility of an *Ising-spin representation*. Indeed, we may rewrite (6.4) as

$$\mathrm{e}^{-V(\phi)} = \sum_{\sigma=\pm1} \mathrm{e}^{-\frac{\kappa}{2}(\phi-\sigma)^2} = C \sum_{\sigma=\pm1} \mathrm{e}^{-\frac{\kappa}{2}\phi_x - \kappa\phi_x\sigma_x} \tag{6.5}$$

where $C := \mathrm{e}^{-\kappa}$. A product of such terms is thus proportional to

$$\prod_x \mathrm{e}^{-V(\phi_x)} \propto \sum_{(\sigma_x)} \prod_x \mathrm{e}^{-\frac{\kappa}{2}\phi_x^2 - \kappa\phi_x\sigma_x} \tag{6.6}$$

This means we can write the Gibbs weight of the model as follows

$$\begin{aligned} &\mathrm{e}^{-\beta\sum_{\langle x,y\rangle}(\phi_x-\phi_y)^2 - \sum_x V(\phi_x)} \\ &\qquad \propto \sum_{(\sigma_x)} \mathrm{e}^{-\beta\sum_{\langle x,y\rangle}(\phi_x-\phi_y)^2 - \frac{\kappa}{2}\sum_x \phi_x^2}\, \mathrm{e}^{-\kappa\sum_x \phi_x\sigma_x} \end{aligned} \tag{6.7}$$

If we elevate (σ_x) to genuine degrees of freedom, we get a model on spins $S_x := (\phi_x, \sigma_x)$ with *a priori* law Lebesgue on $\mathbb{R} \times$ counting measure on $\{-1, 1\}$ and the Hamiltonian

$$\beta H(\phi,\sigma) := \beta \sum_{\langle x,y\rangle} (\phi_x - \phi_y)^2 + \frac{\kappa}{2} \sum_x \phi_x^2 + \kappa \sum_x \phi_x\sigma_x \tag{6.8}$$

Notice the first two terms on the right-hand side is the Hamiltonian of the massive (centered) Gaussian free field while the interaction between the ϕ's and the σ's has on-site form.

Here are some observations whose (simple) proof we leave to the reader:

Lemma 6.1. *Let μ be a Gibbs measure for Hamiltonian* (6.8) *and let ν be its ϕ-marginal. Then ν is a Gibbs measure for the Hamiltonian* (6.1) *subject to* (6.4)*. The marginal ν completely determines μ: For any f depending only on ϕ and σ in a finite set Λ,*

$$E_\mu(f) = E_\nu\Bigl(\sum_{(\sigma_x)_{x\in\Lambda}} f(\phi,\sigma) \prod_{x\in\Lambda} \mathrm{e}^{V(\phi_x) - \frac{\kappa}{2}(\phi_x-\sigma_x)^2} \Bigr) \tag{6.9}$$

We will use $\mathfrak{G}_{\beta,\kappa}$ to denote the set of all Gibbs measures for the Hamiltonian (6.8) with parameters β and κ. The principal result for this model is as follows:

Theorem 6.2. *Let $d \geq 2$. For each $\epsilon > 0$ there is $c > 0$ such that if $\kappa, \kappa/\beta > c$, then there exist $\mu^+, \mu^- \in \mathfrak{G}_{\beta,\kappa}$ which are translation invariant and obey*

$$\mu^{\pm}(\sigma_x = \pm 1) \geq 1 - \epsilon \tag{6.10}$$

and

$$E_{\mu^{\pm}}\big((\phi_x \mp 1)^2\big) \leq \epsilon \tag{6.11}$$

In simple terms, at low temperatures and large curvature of the wells of V, the fields prefer to localize in one of the wells. We remark that, while we chose the model as simple as possible, a similar conclusion would follow for V given by

$$\mathrm{e}^{-V(\phi)} = \mathrm{e}^{-\frac{\kappa_+}{2}(\phi-1)^2+h} + \mathrm{e}^{-\frac{\kappa_-}{2}(\phi+1)^2-h}. \tag{6.12}$$

where h changes the relative weight of the two minima. Indeed, there exists h_{t} at which one has two Gibbs measure — the analogues of μ^+ and μ^-. Moreover, if $\kappa_+ \gg \kappa_-$, then $h_{\mathrm{t}} > 0$ because, roughly speaking, the well at -1 offers "more room" for fluctuations.

6.2 Proof of Phase Coexistence

Here we will prove Theorem 6.2. We will focus on $d = 2$; the proof in general dimension is a straightforward, albeit more involved, generalization.

Let us refer to a face of $\mathbb{Z}^2$ as a *plaquette* (i.e., a plaquette is a square of side one with a vertex of $\mathbb{Z}^2$ in each corner). Given a spin configuration (σ_x), we say that a plaquette is *good* if all four spins take the same value, and *bad* otherwise. Let $\mathcal{B}$ denote the event that the plaquette with lower-left corner at the origin is bad.

Since the interaction is that of the GFF with a modified single-spin measure, the torus Gibbs measure is RP. The crux of the proof is to show that bad plaquettes are suppressed. Specifically, we want to show that

$$\mathfrak{z}(\mathcal{B}) \ll 1 \quad \text{once} \quad \beta, \kappa \gg 1 \tag{6.13}$$

Appealing to the subadditivity lemma (Lemma 5.9) we only need to estimate the $\mathfrak{z}$-value of all possible configurations on the plaquette that

constitute $\mathcal{B}$. Due to the plus-minus symmetry of the σ's, it suffices to examine three patterns:

$$\begin{matrix} - & + \\ + & - \end{matrix} \qquad \begin{matrix} + & - \\ + & - \end{matrix} \qquad \begin{matrix} - & + \\ + & + \end{matrix} \tag{6.14}$$

We begin with the most interesting of the three:

Lemma 6.3. *For any $\beta, \kappa > 0$,*

$$\mathfrak{z}\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right) \le \mathrm{e}^{-\frac{4\beta\kappa}{8\beta+\kappa}} \tag{6.15}$$

Proof. Let $Z_L := \sum_\sigma \int \mathrm{e}^{-\beta H_L(\phi,\sigma)} \prod_{x\in\mathbb{T}_L} \mathrm{d}\phi_x$ be the torus partition function. Given a plaquette spin pattern, let $Z_L(\text{pattern})$ denote the same object with σ fixed to the disseminated pattern — the sole element of $\bigcap_{t\in\mathbb{T}_L} \vartheta_t(\text{pattern})$. (We are working with $B = 1$.) By the definition of $\mathfrak{z}$ we have

$$\mathfrak{z}_L\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)^{|\mathbb{T}_L|} := \frac{Z_L\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)}{Z_L} \le \frac{Z_L\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)}{Z_L\left(\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}\right)} \tag{6.16}$$

Now the partition function with all σ's restricted to $+$ is given by

$$\begin{aligned} Z_L\left(\begin{smallmatrix} + & + \\ + & + \end{smallmatrix}\right) &= \int \mathrm{e}^{-\beta\sum_{\langle x,y\rangle}(\phi_x-\phi_y)^2 - \frac{\kappa}{2}\sum_x \phi_x^2}\, \mathrm{e}^{-\kappa\sum_x \phi_x} \prod_{x\in\mathbb{T}_L} \mathrm{d}\phi_x \\ &= \Big(\dots\Big) E_{\mathrm{GFF}}\big(\mathrm{e}^{-\kappa\sum_x \phi_x}\big) \end{aligned} \tag{6.17}$$

where the expectation is with respect to the massive Gaussian free field and the prefactor denotes the integral of the Gaussian kernel over all ϕ_x. Similarly we obtain

$$Z_L\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right) = \Big(\dots\Big) E_{\mathrm{GFF}}\big(\mathrm{e}^{-\kappa\sum_x \phi_x(-1)^{|x|}}\big) \tag{6.18}$$

where we noticed that by disseminating the pattern $\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}$ we obtain a configuration which is one at even parity x and minus on at odd parity x. Thus we conclude

$$\mathfrak{z}_L\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)^{|\mathbb{T}_L|} \le \frac{E_{\mathrm{GFF}}\big(\mathrm{e}^{-\kappa\sum_x \phi_x}\big)}{E_{\mathrm{GFF}}\big(\mathrm{e}^{-\kappa\sum_x \phi_x(-1)^{|x|}}\big)} \tag{6.19}$$

i.e., we only need to compute the ratio of the Gaussian expectations, and not the prefactors.

Next we recall a standard formula for Gaussian moment generating functions: If X is a multivariate Gaussian, then

$$E(\mathrm{e}^{\lambda\cdot X}) = \mathrm{e}^{\lambda\cdot EX+\frac{1}{2}\mathrm{Var}(\lambda\cdot X)} \tag{6.20}$$

Since $E_{\mathrm{GFF}}(\phi_x) = 0$, we only need to compute the (diagonal) matrix element of $\mathrm{Cov}(\phi_x, \phi_y)$ against vectors $1 = (1, 1, \dots)$ and $(-1)^{|x|}$. However, a quick look at (6.2) will convince us that these functions are eigenvectors of the covariance matrix corresponding to $k = 0$ and $k = (\pi, \pi)$, respectively. Since $\widehat{D}(0) = 0$ while $\widehat{D}(\pi, \pi) = 8$, we get

$$\mathrm{Var}_{\mathrm{GFF}}\Big(\sum_x \phi_x\Big) = \frac{|\mathbb{T}_L|}{\kappa} \tag{6.21}$$

$$\mathrm{Var}_{\mathrm{GFF}}\Big(\sum_x \phi_x(-1)^{|x|}\Big) = \frac{|\mathbb{T}_L|}{8\beta+\kappa} \tag{6.22}$$

where the factor $|\mathbb{T}_L|$ is the (square of) the $L^2(\mathbb{T}_L)$-norm of the functions under consideration. Plugging this in (6.19) we conclude

$$\mathfrak{Z}_L\left(\begin{smallmatrix}-&+\\+&-\end{smallmatrix}\right)^{|\mathbb{T}_L|} \le \exp\left\{\frac{1}{2}|\mathbb{T}_L|\kappa^2\Big(\frac{1}{8\beta+\kappa} - \frac{1}{\kappa}\Big)\right\} \tag{6.23}$$

from which the claim readily follows. □

Next we attend to the other patterns:

Lemma 6.4. *For any $\beta, \kappa > 0$,*

$$\mathfrak{Z}\left(\begin{smallmatrix}+&-\\+&-\end{smallmatrix}\right) \le \mathrm{e}^{-\frac{2\beta\kappa}{4\beta+\kappa}} \tag{6.24}$$

and

$$\mathfrak{Z}\left(\begin{smallmatrix}-&+\\+&+\end{smallmatrix}\right) \le \mathrm{e}^{-\frac{2\beta\kappa}{8\beta+\kappa}} \tag{6.25}$$

Proof. As for (6.24), dissemination of $\left(\begin{smallmatrix}+&-\\+&-\end{smallmatrix}\right)$ leads to alternating stripes of plusses and minuses, i.e., $\sigma_x = (-1)^{|x_1|}$. Again, this is an eigenvector of the covariance matrix (6.2) with $k = (\pi, 0)$. The corresponding $\widehat{D}$ equals 4 and so

$$\mathfrak{Z}_L\left(\begin{smallmatrix}+&-\\+&-\end{smallmatrix}\right)^{|\mathbb{T}_L|} \le \exp\left\{\frac{1}{2}|\mathbb{T}_L|\kappa^2\Big(\frac{1}{4\beta+\kappa} - \frac{1}{\kappa}\Big)\right\} \tag{6.26}$$

yielding (6.24).

The pattern $\left(\begin{smallmatrix}-&+\\+&+\end{smallmatrix}\right)$ is more complex because its dissemination will not lead to an eigenvector of the covariance matrix. However, we circumvent this problem by noting that Lemma 5.10 implies

$$\mathfrak{Z}\left(\begin{smallmatrix}-&+\\+&+\end{smallmatrix}\right) \le \mathfrak{Z}\left(\begin{smallmatrix}-&+\\+&-\end{smallmatrix}\right)^{1/2}\mathfrak{Z}\left(\begin{smallmatrix}+&+\\+&+\end{smallmatrix}\right)^{1/2} \le \mathfrak{Z}\left(\begin{smallmatrix}-&+\\+&-\end{smallmatrix}\right)^{1/2} \tag{6.27}$$

where we used $\mathfrak{Z}\left(\begin{smallmatrix}+&+\\+&+\end{smallmatrix}\right) \le 1$. Now (6.25) follows from Lemma 6.3. □

Corollary 6.5. *For each $\epsilon > 0$ there exists $a > 0$ such that if $\beta, \kappa > a$, then $\mathfrak{z}(\mathcal{B}) \le \epsilon$.*

Proof. The event $\mathcal{B}$ can be written as the union over a finite number of bad patterns. On the basis of Lemmas 6.3–6.4 the claim holds for $\mathcal{B}$ replaced by any fixed bad pattern. The desired bound now follows — with slightly worse constants — by invoking Lemma 5.9. □

Next we explain our focus on the bad event:

Lemma 6.6. *There exists a constant $c \in (1, \infty)$ such that if $c\mathfrak{z}(\mathcal{B}) < 1/2$ then for any $x, y \in \mathbb{T}_L$,*

$$\mu_L(\sigma_x = 1,\, \sigma_y = -1) \le 2c\mathfrak{z}(\mathcal{B}). \tag{6.28}$$

Proof. This is a consequence of a simple Peierls' estimate. Indeed, if $\sigma_x = 1$ and $\sigma_y = -1$, then x is separated from y by a "circuit" of bad plaquettes. (Formally, either all plaquettes containing x are bad or there exists a non-trivial connected component of good — i.e., *not* bad — plaquettes containing x. This component cannot cover the whole torus because $\sigma_y = -1$; the above "circuit" is then comprised of the bad plaquettes on the boundary of this component.) This means that

$$\mu_L(\sigma_x = 1,\, \sigma_y = -1) \le \sum_\gamma \mu_L\Bigl(\bigcap_{t\in\gamma} \vartheta_t(\mathcal{B})\Bigr) \le \sum_\gamma \mathfrak{z}(\mathcal{B})^{|\gamma|} \tag{6.29}$$

where $|\gamma|$ denotes the maximal number of disjoint bad plaquettes in γ and where we used the chessboard estimates to derive the second bound. By standard arguments, the number of circuits of "length" n surrounding x or winding around $\mathbb{T}_L$ at least once is bounded by c^n, for some constant $c > 1$. It follows

$$\mu_L(\sigma_x = 1,\, \sigma_y = -1) \le \sum_{n\ge 1} c^n \mathfrak{z}(\mathcal{B})^n \tag{6.30}$$

Under the condition $c\mathfrak{z}(\mathcal{B}) < 1/2$ this sum is less than twice its first term. □

Finally, we can assemble the ingredients into the desired proof of phase coexistence:

Proof of Theorem 6.2. By symmetry of the torus measure, we have

$$\mu_L(\sigma_x = 1) = 1/2 = \mu_L(\sigma_x = -1). \tag{6.31}$$

Let z be a site at the back of the torus — that is distant at least $L/2$ from the origin — and define

$$\mu_L^{\pm}(-) := \mu_L(-|\sigma_z = \pm 1). \tag{6.32}$$

These measures satisfy the DLR condition with respect to any function that depends only on the "front" of the torus and so any weak cluster point of these measures will be an infinite-volume Gibbs measure. Extract such measures by subsequential limits and call them μ^+ and μ^-, respectively.

We claim that $\mu^+ \neq \mu^-$. Indeed, by Lemma 6.6 we have

$$\mu_L^+(\sigma_x = -1) \leq 2c\mathfrak{z}(\mathcal{B}) \tag{6.33}$$

once $\mathfrak{z}(\mathcal{B}) \ll 1$ and, by Corollary 6.5, this actually happens once $\beta, \kappa \gg 1$. Thus if, say, $2c\mathfrak{z}(\mathcal{B}) \leq 1/4$, then $\mu_L^+(\sigma_x = -1) \leq 1/4$ and, at the same time, $\mu_L^-(\sigma_x = +1) \leq 1/4$. The same holds for the limiting objects and so $\mu^+ \neq \mu^-$. Note that the measures can be averaged over shifts so that they become translation invariant. □

Notice that in the last step of the proof we used, rather conveniently, the plus-minus symmetry of the torus measure. In the asymmetric cases, e.g., (6.12), one can either invoke a continuity argument — choose $h = h_L$ such that (6.31) holds — or turn (6.28) into the proof that $|\mathbb{T}_L|^{-1} \sum_{x \in \mathbb{T}_L} \sigma_x$ will take values in $[-1, -1 + \epsilon] \cup [1 - \epsilon, 1]$ with probability tending to one as $L \to \infty$. The latter "forbidden-gap" argument is rather robust and extends, with appropriate modifications, to all shift-ergodic infinite-volume Gibbs measures. Hence, the empirical magnetization in ergodic measures cannot change continuously with h.

To prove Theorem 6.2, it remains to show the concentration of the ϕ's around the σ's:

Proof of (6.11). Let μ be a Gibbs measure for parameters β and κ. Then (6.8) shows that, conditional on the σ's, the ϕ's are Gaussian with mean

$$E_\mu(\phi_x|\sigma) = \kappa\big((2\beta\Delta + \kappa)^{-1}\sigma\big)_x \tag{6.34}$$

and covariance $(2\beta\Delta + \kappa)^{-1}$, where Δ is the lattice Laplacian. Now, once $\beta/\kappa \ll 1$ we may expand the inverse operator into a power series to get

$$E_\mu(\phi_x|\sigma) - \sigma_x = \sum_{n \geq 1} \Big(\frac{2\beta}{\kappa}\Big)^n (\Delta^n \sigma)_x \tag{6.35}$$

which by the fact that $|\sigma_z| = 1$ is $O(\beta/\kappa)$ independently of x. Since the conditional variance of ϕ_x is $O(1/\kappa)$, we obtain

$$\begin{aligned} E_\mu\big((\phi_x - \sigma_x)^2\big|\sigma\big) &\leq 2E_\mu\big([\phi_x - E(\phi_x|\sigma)]^2\big|\sigma\big) + 2\big(E(\phi_x|\sigma) - \sigma_x\big)^2 \\ &= O\big((\beta/\kappa)^2\big) + O(1/\kappa) \end{aligned} \tag{6.36}$$

with the constants implicit in the O's independent of the σ's and x. Thus, if $\kappa \gg 1$ and ${}^{\kappa}/_{\beta} \gg 1$, then (6.11) follows by the fact that $\mu^{\pm}$ put most of the mass on $\sigma_x = \pm 1$. □

6.3 Gradient Fields with Non-convex Potential

Having demonstrated the use of chessboard estimates on a toy model, we will proceed to discuss more complicated systems. We begin with an example which is somewhat similar to the Gaussian double-well model.

A natural generalization of the massless GFF is obtained by replacing the quadratic gradient interaction by a general, even function of the gradients. The relevant Hamiltonian (again with temperature incorporated in it) is

$$\beta H(\phi) := \sum_{\langle x,y \rangle} V(\phi_x - \phi_y) \tag{6.37}$$

The requirements that we generally put on V are continuity, evenness and quadratic growth at infinity. Under these conditions one can always define finite-volume Gibbs measures.

As to the measures in infinite volume, the massless nature of the model may prevent existence of a meaningful thermodynamic limit in low dimensions; however, if one restricts attention to *gradient variables*,

$$\eta_b := \phi_y - \phi_x \quad \text{if } b \text{ is the oriented edge } (x,y), \tag{6.38}$$

then the infinite-volume Gibbs measures exist, and may be characterized by a DLR condition, in all $d \geq 1$. We call these *gradient Gibbs measures* (GGM). A non-trivial feature of the GGM is that they obey a host of constraints. Indeed, almost every η is such that

$$\eta_{b_1} + \eta_{b_2} + \eta_{b_3} + \eta_{b_4} = 0 \tag{6.39}$$

for any plaquette $(b_1, \dots, b_4)$ with bonds listed (and oriented) in the counterclockwise direction.

Surprisingly, the classification of all possible translation-invariant, infinite-volume GGMs can be achieved under the condition that V is strictly convex:

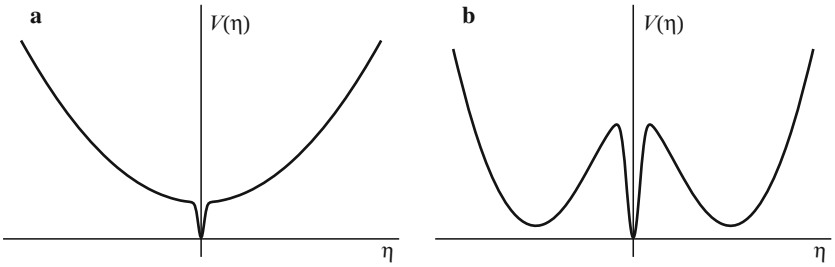
Theorem 6.7. *Suppose V is convex, twice continuously differentiable with V'' bounded away from zero and infinity. Then the shift-ergodic GGMs μ are in one-to-one correspondence with their tilt, which is a vector $a \in \mathbb{R}^d$ such that*

$$E_\mu(\eta_b) = a \cdot b \tag{6.40}$$

for every (oriented) bond b (we regard b as a unit vector for this purpose).

The word *tilt* comes from the interpretation of a as the slope or the incline of the interface whose height-gradient along bond b is given by η_b. The proof of this result — which is due to Funaki and Spohn — is based on the use of the Brascamp-Lieb inequality through which the convexity assumption enters in an essential way. It is also known that the large-scale fluctuation structure of the η's is that of a Gaussian Free Field.

A natural question to ask is what happens when V is *not* convex. Specific examples of interest might be V taking the form of a double-well potential — kind of like for the Gaussian double-well model — or V's as in the figure:



As it turns out, the double-well case is not quite tractable at the moment — and most likely behaves like a massless GFF on large scales — but the other two cases are within reach. We will focus on the case (a) and, as for the Gaussian double-well model, assume a particular form of the potential:

$$\mathrm{e}^{-V(\eta)} := p\,\mathrm{e}^{-\kappa_{\mathrm{O}}\eta^2/2} + (1-p)\,\mathrm{e}^{-\kappa_{\mathrm{D}}\eta^2/2} \tag{6.41}$$

where κ_{O} and κ_{D} are positive numbers and $p \in [0,1]$ is a parameter to be varied. For this system one can prove the following result:

Theorem 6.8. *Suppose $d = 2$ and $\kappa_{\mathrm{O}} \gg \kappa_{\mathrm{D}}$. Then there is $p_{\mathrm{t}} \in (0,1)$ and, for V with $p = p_{\mathrm{t}}$, there are two distinct, infinite-volume, shift-ergodic GGMs μ_{ord} and μ_{dis} that are invariant with respect to lattice rotations and have the following properties:*

(1) zero tilt:

$$\frac{1}{|\Lambda_L|} \sum_{\substack{b=(x,y)\\ x,y\in\Lambda_L}} \eta_b \underset{L\to\infty}{\longrightarrow} 0, \qquad \mu_{\mathrm{ord}}, \mu_{\mathrm{dis}}\text{-a.s.} \tag{6.42}$$

(2) distinct fluctuation size:

$$E_{\mu_{\text{ord}}}(\eta_b^2) \ll E_{\mu_{\text{dis}}}(\eta_b^2) \tag{6.43}$$

The upshot of this result is that, once the convexity of V is strongly violated, the conclusions of Theorem 6.7 do not apply. While the example is restricted to $d = 2$, and to potentials of the form (6.41), generalizations to $d \geq 2$ and other potentials as in the above figure are possible and reasonably straightforward.

Here are the main steps of the proof. First, as for the Gaussian double-well model, we use (6.41) to expand the Gibbs weight according to whether the first or the second term in (6.41) applies. This gives rise to a configuration of coupling strengths (κ_b), one for each bond b, which take values in $\{\kappa_{\text{O}}, \kappa_{\text{D}}\}$. The joint Hamiltonian of the η's and the κ's is

$$\beta H(\eta, \kappa) := \sum_b \frac{\kappa_b}{2} \eta_b^2 \tag{6.44}$$

The joint measure is RP with respect to reflections through bonds and sites and, conditional on (κ_b), the η's are Gaussian.

For the proof of phase coexistence, we focus on lattice plaquettes and divide these into good and bad according to whether all of the edges have the same coupling κ or not. The dissemination of each bad patterns leads to a Gaussian integral but this time for GFF with inhomogeneous — yet periodically varying — couplings. For instance, the pattern with three bonds of type κ_{O} and one of type κ_{D} disseminates into periodic configuration where the edges on every other vertical line is of type κ_{D} and all other edges are of type κ_{O}. Similarly for all other bad patterns.

The periodic nature of the disseminated events allows the use of Fourier modes — i.e., pass to the reciprocal torus — to diagonalize the requisite covariance matrices. For instance, the aforementioned pattern with three κ_{O}'s and one κ_{D} leads to a configuration which is periodic with period two. A calculation shows that the covariance is block diagonal with 2×2 blocks of the form

$$\Pi(k) := \begin{pmatrix} \kappa_{\text{O}}|a_-|^2 + \frac{1}{2}(\kappa_{\text{O}} + \kappa_{\text{D}})|b_-|^2 & \frac{1}{2}(\kappa_{\text{O}} - \kappa_{\text{D}})|b_-|^2 \\ \frac{1}{2}(\kappa_{\text{O}} - \kappa_{\text{D}})|b_-|^2 & \kappa_{\text{O}}|a_+|^2 + \frac{1}{2}(\kappa_{\text{O}} + \kappa_{\text{D}})|b_-|^2 \end{pmatrix} \tag{6.45}$$

where $a_\pm$ and $b_\pm$ are defined by

$$a_\pm = 1 \pm \mathrm{e}^{\mathrm{i}k_1} \quad \text{and} \quad b_\pm = 1 \pm \mathrm{e}^{\mathrm{i}k_2} \tag{6.46}$$

with $k := (k_1, k_2)$ varying through one half of the reciprocal torus $\mathbb{T}_L^*$. (The block combines the contribution of both k and $k + \pi\hat{e}_1$, and so we only need half of all k's.) The requisite Gaussian integral then reduces to $\prod_{k\in\mathbb{T}_L^*\setminus\{0\}}[\det \Pi(k)]^{-1/4}$ where in the exponent we get $1/4$ instead of the expected $1/2$ to account for double counting of the k's. To estimate the growth rate of this product, we note that

$$\prod_{k\in\mathbb{T}_L^*\setminus\{0\}} [\det \Pi(k)]^{-1/4} = \exp\left\{-|\mathbb{T}_L|\frac{1}{4}\int \frac{\mathrm{d}k}{(2\pi)^2}\log\det \Pi(k) + o(|\mathbb{T}_L|)\right\} \tag{6.47}$$

The integral plays the role of the free energy associated with the Gaussian variables on the background of the specific periodic configuration of the κ's. A similar expression — with different integrand — applies to each pattern.

Comparing the integrals for all possible arrangements of the two types of bonds around a plaquette, we find that under the condition $\kappa_{\mathrm{O}} \gg \kappa_{\mathrm{D}}$, the bad patterns are heavily suppressed. Thus bad plaquettes are infrequent and can be regarded as parts of a contour. As it is not possible to pass from all-κ_{O} pattern to all-κ_{D} pattern without crossing a bad plaquette, the coexistence follows — as for the double-well model — by a standard Peierls' argument and chessboard estimates. Full details of the proof are to be found in a paper by Kotecký and the present author.

The two-dimensional model has the special feature that we can actually compute p_{t}:

Theorem 6.9. *Let $d = 2$. If $\kappa_{\mathrm{O}}/\kappa_{\mathrm{D}} \gg 1$, then p_{t} is given by*

$$\frac{p_{\mathrm{t}}}{1-p_{\mathrm{t}}} = \left(\frac{\kappa_{\mathrm{D}}}{\kappa_{\mathrm{O}}}\right)^{1/4}. \tag{6.48}$$

This is a consequence of a duality relation that can be used to exchange the roles of κ_{O} and κ_{D}. It is also interesting to note that, while the one-to-one correspondence between the Gibbs measures and their tilt is violated for non-convex potentials, the large-scale fluctuation structure remains that of a Gaussian Free Field. Indeed, we have:

Theorem 6.10. *Let $d = 2$. For each translation-invariant, ergodic gradient Gibbs measure μ with zero tilt, there exists a positive-definite $d\times d$ matrix $q = q(\mu)$ such that for any smooth $f\colon \mathbb{R}^2 \to \mathbb{R}$ with compact support and $\int f(x)\mathrm{d}x = 0$,*

$$\int \mathrm{d}x\, \phi_{\lfloor x/\epsilon \rfloor} f(x) \;\xrightarrow[\epsilon\downarrow 0]{\mathcal{D}}\; \mathcal{N}\bigl(0, (f, Q^{-1} f)\bigr) \tag{6.49}$$

where $\mathcal{N}(0,C)$ denotes a normal random variable with mean zero and covariance C and Q is the elliptic operator

$$Qf(x) := \sum_{i,j=1}^{d} q_{ij} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \tag{6.50}$$

The basis of this result — derived in all $d \geq 1$ by Spohn and the present author — is the fact that, conditional on the κ's, the ϕ's are Gaussian with mean zero and covariance given by the inverse of the generator of a reversible random walk in random environment. The Gaussian limit is a consequence of an (annealed) invariance principle for such random walks and some basic arguments in homogenization theory. The restriction to zero tilt appears crucially in the proof.

6.4 Spin-waves vs Infinite Ground-state Degeneracy

Next we will discuss a couple of spin models whose distinctive feature is a high degeneracy of their ground state which is removed, at positive temperature, by soft-mode spin-wave fluctuations. The simplest example with such property is as follows:

Orbital compass model: Here the spins on $\mathbb{Z}^d$ take values in a unit sphere in $\mathbb{R}^d$, i.e., $S_x \in \mathbb{S}^{d-1}$ with $x \in \mathbb{Z}^d$. The Hamiltonian is

$$H(S) := \sum_{x} \sum_{\alpha=1}^{d} (S_x^{(\alpha)} - S_{x+\hat{\mathrm{e}}_\alpha}^{(\alpha)})^2 \tag{6.51}$$

where $S_x^{(\alpha)}$ denotes the α-th Cartesian component of the spin and $\hat{\mathrm{e}}_\alpha$ is the unit vector in the α-th coordinate direction.

Despite a formal similarity with the Heisenberg model, note that only one component of the spin is coupled in each lattice direction. Notwithstanding, *every* constant configuration is still a minimum-energy state of (6.51). Further ground states may be obtained from the constant ones by picking a coordinate direction α and changing the sign of the α-th component of all spins in some of the "lines" parallel with $\hat{\mathrm{e}}_\alpha$. In $d = 2$ these are all ground states but in $d \geq 3$ other operations are possible that preserve the minimum-energy property.

The key question is now what happens with this huge ground-state degeneracy at positive temperatures. Here is a theorem one can prove about the two-dimensional system:

Theorem 6.11. *For each $\epsilon > 0$ there exist $\beta_0 > 0$ and, for each $\beta \geq \beta_0$, there exist two distinct, shift-ergodic Gibbs measures $\mu_1, \mu_2 \in \mathfrak{G}_\beta$ such that*

$$E_{\mu_j}\bigl(|S_x \cdot \hat{\mathrm{e}}_j|\bigr) \geq 1 - \epsilon, \qquad j = 1, 2 \tag{6.52}$$

Moreover, for any $\mu \in \mathfrak{G}_\beta$ we have

$$E_\mu(S_x) = 0 \tag{6.53}$$

and there are no shift-ergodic $\mu \in \mathfrak{G}_\beta$, $\beta \geq \beta_0$, for which we would have $\max_{j=1,2} E_\mu\bigl(|S_x \cdot \hat{\mathrm{e}}_j|\bigr) < 1 - \epsilon$.

The main idea underlying the proof is the evaluation of the free energy associated with spin-wave perturbations of the constant ground states; it this expected that only the states with the largest contribution of these fluctuations survive at positive temperatures. Specifically, we need to quantify the growth rate of the torus partition function with all spins constrained to lie within Δ of a given direction:

Lemma 6.12. *For each $\epsilon > 0$ there is $\delta > 0$ such that if β, Δ obey*

$$\beta\Delta^2 > \frac{1}{\delta} \qquad \text{and} \qquad \beta\Delta^3 < \delta \tag{6.54}$$

then for every $\hat{\mathrm{v}}_\theta := (\cos\theta, \sin\theta) \in \mathbb{S}^1$,

$$E_{\otimes\mu_0}\Bigl(\mathrm{e}^{-\beta H_L(S)} \prod_{x \in \mathbb{T}_L} 1_{\{|S_x - \hat{\mathrm{v}}_\theta| < \Delta\}}\Bigr) = \Bigl(\frac{2\pi}{\beta}\Bigr)^{L^2/2} \mathrm{e}^{-L^2[F(\theta) + o(\epsilon)]} \tag{6.55}$$

where

$$F(\theta) := \frac{1}{2} \int \frac{\mathrm{d}k}{(2\pi)^2} \log\bigl\{\sin^2(\theta)|1 - \mathrm{e}^{\mathrm{i}k_1}|^2 + \cos^2(\theta)|1 - \mathrm{e}^{\mathrm{i}k_2}|^2\bigr\} \tag{6.56}$$

The quantity F has the interpretation of the *spin-wave free energy* where the term "spin wave" refers to slowly varying deformations of a constant ground states. A convexity argument — based on the identity $\sin^2(\theta) + \cos^2(\theta) = 1$ — now shows that F is minimized by $\theta = 0, \pi/2, \pi, 3\pi/2$, i.e., exactly in one of the coordinate directions. This corroborates the intuition that only the configurations with most of the spins aligned in one of these directions will be relevant at low temperatures. However, to extract a proof of phase coexistence, we will have to again invoke a Peierls' argument.

Fix $\kappa > 0$ and let $\Delta := \beta^{-\frac{5}{12}}$ and $B := \log\beta$ and let $\mathcal{B}_{\text{E}}$ and $\mathcal{B}_{\text{SW}}$ denote the following events:

(1) $\mathcal{B}_{\text{E}} := \{$ a pair of neighboring spins in Λ_B differ by an angle $\geq \Delta\}$
(2) $\mathcal{B}_{\text{SW}}$ is the set of configurations in the complement of $\mathcal{B}_{\text{E}}$ in which the block Λ_B has all neighboring spins within Δ of each other with at least $\kappa \gg \Delta$ from one of the four coordinate directions

The event $\mathcal{B}_{\text{E}}$ captures the situations when two neighboring spins are not quite close to each other leading to excess energy order Δ^2. As a result of that,

$$\mathfrak{z}(\mathcal{B}_{\text{E}}) \leq 3B^3 \text{e}^{-c_3\beta\Delta^2} \tag{6.57}$$

The event $\mathcal{B}_{\text{SW}}$ collects the configurations where the energy is good but the fluctuations are not sufficiently powerful. The calculation in Lemma 6.12 and a simple use of the subadditivity lemma show

$$\mathfrak{z}(\mathcal{B}_{\text{SW}}) \leq \frac{c_1}{\Delta} \text{e}^{-c_2 B^3 \kappa^2} \tag{6.58}$$

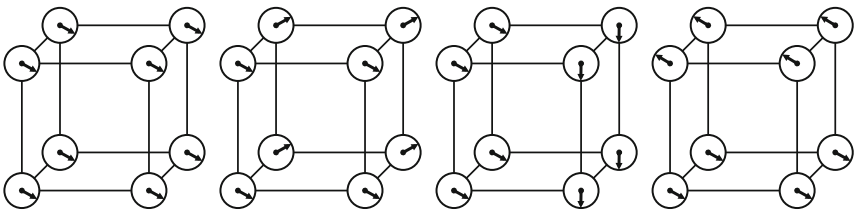
for some constants $c_1, c_2 > 0$. Thus, for our choices of Δ and B, once $\beta \gg 1$ the density of blocks where $\mathcal{B}_{\text{E}} \cup \mathcal{B}_{\text{SW}}$ occurs in any typical configuration from the torus measure will be rather small. However, if a block is aligned in one coordinate direction and another block is aligned in a different direction, they must be separated by a "circuit" of bad blocks. Such circuits are improbable which leads to phase separation. Details of these calculations — which extend even to quantum setting — can be found in a paper by Chayes, Starr and the present author.

120-degree model: A somewhat more complicated version of the interaction, but with the spins S_x taking values in the unit circle $\mathbb{S}^1$, can be contrived in $d = 3$. The Hamiltonian will actually look just as for the orbital compass model except that $S_x^{(\alpha)}$ are not Cartesian components but projections on the three third-roots of unity $\hat{\text{b}}_1, \hat{\text{b}}_2, \hat{\text{b}}_3$ in $\mathbb{S}^1$. Explicitly,

$$H(S) := \sum_x \sum_{\alpha=1,2,3} \bigl(S_x \cdot \hat{\text{b}}_\alpha - S_{x+\hat{\text{e}}_\alpha} \cdot \hat{\text{b}}_\alpha\bigr)^2 \tag{6.59}$$

Again, all constant configurations are ground states and further ground states may again be obtained by judicious reflections. Fortunately, the number of energy-preserving operations one can perform on ground states is much smaller than for the orbital compass model, and all

ground states can thus be classified. Namely, given a ground state configuration, every unit cube in $\mathbb{Z}^3$ looks as one of the four cubes in the picture



modulo, of course, a simultaneous rotation of all spins. Here is what we can we say rigorously about this model:

Theorem 6.13. *Let $\hat{\mathrm{w}}_1, \ldots, \hat{\mathrm{w}}_6 \in \mathbb{S}^1$ be the six sixth roots of unity. For each $\epsilon > 0$ there exist $\beta_0 > 0$ and, for each $\beta \geq \beta_0$, there exist six distinct, shift-ergodic Gibbs measures $\mu_1, \ldots, \mu_6 \in \mathfrak{G}_\beta$ such that*

$$E_{\mu_j}\big(S_x \cdot \hat{\mathrm{w}}_j\big) \geq 1 - \epsilon, \qquad j = 1, \ldots, 6 \tag{6.60}$$

There are no shift-ergodic $\mu \in \mathfrak{G}_\beta$, $\beta \geq \beta_0$, for which we would have $\max_{j=1,\ldots,6} E_\mu\big(S_x \cdot \hat{\mathrm{w}}_j\big) < 1 - \epsilon$.

The ideas underlying this theorem are quite similar to the orbital compass model. First we find out that the spin-wave free energy for fluctuations about the ground state pointing in direction θ is given by

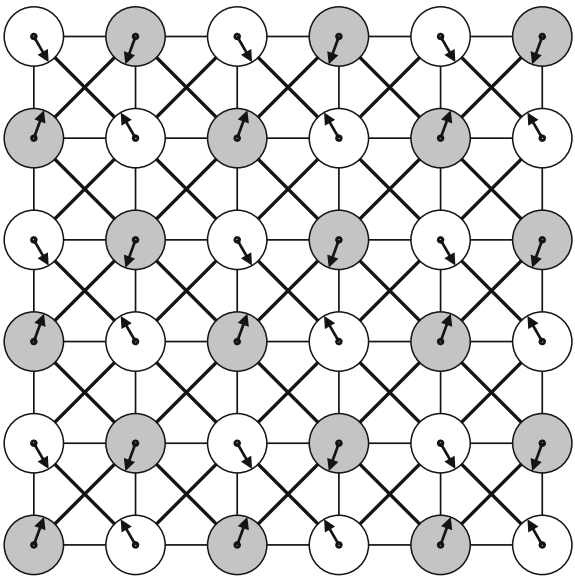
$$F(\theta) := \frac{1}{2} \int \frac{\mathrm{d}k}{(2\pi)^3} \left[\log \sum_{\alpha=1,2,3} q_\alpha(\theta) |1 - \mathrm{e}^{\mathrm{i}k_\alpha}|^2 \right] \tag{6.61}$$

where $q_1 := \sin^2(\theta)$, $q_2 := \sin^2(\theta - 120^\circ)$ and $q_3 := \sin^2(\theta + 120^\circ)$. A surprisingly sophisticated argument is then required to show that F is minimal only for θ of the form $\frac{\pi}{3} j$, $j = 1, \ldots, 6$. Once we have this information, the rest of the argument follows a route very similar to that for the orbital compass model (including the introduction of the scales κ and Δ and the corresponding events $\mathcal{B}_{\mathrm{E}}$ and $\mathcal{B}_{\mathrm{SW}}$). Details appeared in a paper by Chayes, Nussinov and the present author.

n.n. and n.n.n. antiferromagnet: Finally, we will consider a toy model that exemplifies the features of both systems above. Here $d = 2$ and the spins take again values in $\mathbb{S}^1$, but the interaction is *antiferromagnetic* — that is, with a preference for antialignment — for both nearest and next-nearest neighbors:

$$H(S) := \gamma \sum_x \bigl[S_x \cdot S_{x+\hat{e}_1} + S_x \cdot S_{x+\hat{e}_2}\bigr] + \sum_x \bigl[S_x \cdot S_{x+\hat{e}_1+\hat{e}_2} + S_x \cdot S_{x+\hat{e}_1-\hat{e}_2}\bigr] \tag{6.62}$$

Assuming $|\gamma| < 2$, the minimum energy state is obtained by first enforcing the n.n.n. constraints — there is an antiferromagnetic, or Neél, order on both even and odd sublattice — and only then worrying about how to satiate the n.n. constraint. But once the sublattices are ordered antiferromagnetically, the net interaction between the sublattices is zero — and so each of the sublattices can be rotated independently! Here is a configuration of this form:



For this system we can nevertheless prove the following theorem:

Theorem 6.14. *For each $\epsilon > 0$ there exist $\beta_0 > 0$ and, for each $\beta \geq \beta_0$, there exist two distinct, shift-ergodic Gibbs measures $\mu_1, \mu_2 \in \mathfrak{G}_\beta$ such that*

$$-E_{\mu_j}\bigl(S_x \cdot S_{x+\hat{e}_1 \pm \hat{e}_2}\bigr) \geq 1 - \epsilon \tag{6.63}$$

and

$$E_{\mu_j}\bigl(S_x \cdot S_{x+\hat{e}_j}\bigr) \geq 1 - \epsilon, \qquad j = 1, 2 \tag{6.64}$$

There are no shift-ergodic $\mu \in \mathfrak{G}_\beta$, $\beta \geq \beta_0$, for which either (6.63) *or at least one of* (6.64) *does not hold.*

As for the two models above, everything boils down to a spin-wave calculation. Here the relevant parameter is the relative orientation θ of the two antiferromagnetically ordered sublattices. The spin-wave free energy is then

$$F(\theta) := \frac{1}{2} \int_{[-\pi,\pi]^2} \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2} \log D_{\boldsymbol{k}}(\theta) \tag{6.65}$$

where

$$\begin{aligned} D_{\boldsymbol{k}}(\theta) := {} & |1 - \mathrm{e}^{\mathrm{i}(k_1+k_2)}|^2 + |1 - \mathrm{e}^{\mathrm{i}(k_1-k_2)}|^2 \\ & + \gamma \cos(\theta) \big(|1 - \mathrm{e}^{\mathrm{i}k_1}|^2 - |1 - \mathrm{e}^{\mathrm{i}k_1}|^2 \big) \end{aligned} \tag{6.66}$$

As $D(\theta) = \alpha D(0) + (1-\alpha) D(\pi)$, with $\alpha := \frac{1}{2}(1 + \cos(\theta))$, Jensen's inequality for the logarithm directly shows that F is minimized by $\theta = 0$ or $\theta = \pi$. In spin configurations, the former corresponds to horizontal alignment and vertical antialignment of nearest neighbors, and the latter to horizontal antialignment and vertical alignment, i.e., stripe states. Details of all calculations appeared in a paper by Chayes, Kivelson and the present author.

Notice that, despite the fact that the lattices maintain a specific *relative* orientation at low temperatures, a Mermin-Wagner argument ensures that every Gibbs measure is invariant under a rigid rotation of all spins.

6.5 Literature Remarks

The Gaussian double-well model is a standard example which can be treated either by methods of reflection positivity, or by Pirogov-Sinai theory [35]. Representations of the kind (6.4) have been used already before, e.g., by Külske [74, 75] and Zahradník [107]. The method of proof presented here draws on the work of Dobrushin, Kotecký and Shlosman [33, 71, 69] which was used to control order-disorder transitions in a number of systems; most notably, the q-state Potts model with $q \gg 1$ [69]. These methods may be combined with graphical representations of Edwards-Sokal [39] (or Fortuin-Kasteleyn [43]) to establish rather complicated phase diagrams, e.g., [26, 12]. Recently, the method has been used to resolve a controversy about a transition can occur in 2D non-linear vector models [41, 42].

Theorem 6.7 has been proved by Funaki and Spohn [53]. As already mentioned, their proof is based on convexity properties of the potential V — by invoking the Brascamp-Lieb inequality as well as certain

coupling argument to the natural dynamical version of the model — and so it does not extend beyond the convex case. (A review of the gradient measures, and further intriguing results, can be found in Funaki [52], Velenik [104] or Sheffield [92].) Theorem 6.8 was proved by Biskup and Kotecký [17]; Theorem 6.10 was derived by Biskup and Spohn [18].

The interest in models in Sect. 6.4 came from a physics controversy about whether orbital ordering in transition-metal oxides exists at low temperatures. On the basis of rigorous work by Biskup, Chayes and Nussinov [13] (120-degree model) Biskup, Chayes, Nussinov and Starr [15, 14] (2D and 3D orbital compass model), it was demonstrated that, at least at the level of classical models, spin-wave fluctuations stabilize certain ground states [83]. The conclusions hold also the 2D quantum orbital-compass model with large quantum spins [15]. The mechanism of entropic stabilization — or, in physics jargon, *order by disorder* — is most clearly demonstrated in the n.n. & n.n.n. antiferromagnet studied by Biskup, Chayes and Kivelson [11]. This model actually goes back to the papers by Shender [93] and Henley [63] which first spelled out the original order-by-disorder physics arguments.

All three "phase coexistance" theorems in Sect. 6.4 have, apart from an existence clause, also a clause on the *absence* of ergodic states whose local properties deviate from those whose existence was asserted. Actually, these were not the content of the original work [13, 11] because, at that time, the focus on torus measures dictated by reflection positivity was deemed to make it impossible to *rule out* the occurrence of some exotic measures. A passage to such statements was opened by the work of Biskup and Kotecký [16]; the non-existence clauses in Theorems 6.11, 6.13 and 6.14 are direct consequences of the main result of [16] and the method of proof of the existence part. This technique does not quite apply in the setting of gradient models due to the strong role the boundary conditions play in this case.

7 Topics Not Covered

There are naturally many interesting topics dealing with reflection positivity that have not been covered by these notes. Here we will attempt to at least provide a few relevant comments and give pointers to the literature where an interested reader may explore the subject to the desired level of detail.

The first (and large) area which was neglected is that of *quantum* models. Here one faces the principal difficulty that the spin variables

are replaced by operators which, generally, do not commute with one another. Nevertheless, reflection positivity can be proved for reflections through planes between sites under the condition that the Hamiltonian is of the form (5.10). (For reflections through planes of sites the non-commutativity of involved objects makes the above technology largely unavailable.) Thus, chessboard estimates and, by a passage via the Duhamel two-point function, also infrared bound can again be established. This and the resulting applications to proofs of phase transitions in, e.g., the quantum Heisenberg *anti*ferromagnet and XY-model constitute the papers of Dyson, Lieb and Simon [38] and Fröhlich and Lieb [48]. A pedagogical account of these can be found in the notes by Tóth [103].

Unlike for the classical models, in the quantum setting reflection positivity appears to be a somewhat peculiar condition. Generally, it requires that the involved operators can be represented by either real or purely imaginary matrices. This is where the technique fails in the case of the quantum Heisenberg ferromagnet (Speer [99], but see also Kennedy [67] and Conlon and Solovej [27]). Notwithstanding, the technique continued to be applied in the quantum world to derive useful conclusions; e.g., to study long range order in two-dimensional antiferromagnets (Kennedy, Lieb and Shastry [68]), to resolve the so called flux phase problem in the Hubbard model (Lieb [78], see also Macris and Nachtergaele [81]) or to prove uniqueness of the ground state in the half-filled band therein (Lieb [77]). The latter work invokes *spin-reflection positivity*; a new idea later further exploited by, e.g., Tian [102] and Tasaki [101]. Other applications of reflection positivity in itinerant-electron models appear in, e.g., Macris [79] and Macris and Lebowitz [80].

As already mentioned, one can use RP to develop a rigorous link between the phase transitions in quantum and classical systems (Biskup, Chayes and Starr [15]). Here the main idea is the conversion of the quantum chessboard estimate to the classical one by means of an extension of Berezin-Lieb inequalitites to matrix elements in the basis of coherent states.

Another topic not sufficiently represented in these notes is that of dimer or other combinatorial models. Here we wish to mention, e.g., the conclusions concerning the six-vertex model and hard-core lattice gasses (Fröhlich, Israel, Lieb and Simon [47]) or the liquid-crystal models based on interacting dimers (Heilmann and Lieb [62] and Abraham and Heilmann [1]). There is also a novel application to characterization of graph homomorphisms (Freedman, Lovász and Schrijver [44]).

The origin of reflection positivity lies within the field theory as part of the Osterwalder-Schrader axioms. A reader interested in this direction should employ the relevant search outlets to explore the literature on the subject. For statistical mechanics, interesting applications come in the proofs of phase transitions in Euclidean field theories, e.g., that of quark confinement (Borgs and Seiler [20]) or chiral symmetry breaking (Salmhofer and Seiler [91]) in gauge theories.

Finally, there is the recent clever application of chessboard estimates to control the rigidity of Dobrushin interfaces in the Ising (and some other) three dimensional models (Shlosman and Vignaud [96]). This direction will likely be further exploited to study interface states in continuum-spin systems.

8 Three Open Problems

We finish with a brief discussion of three general open problems of the subject covered by these notes which the present author finds worthy of significant research effort.

In Chapters 3 and 4 we have shown how useful the infrared bound is in proofs of symmetry breaking and control of the mean-field approximation. Unfortunately, the only way we currently have for proving the IRB is reflection positivity. So our first problem is:

Problem 8.1. Consider models with the Hamiltonian $H = -\sum_{\langle x,y\rangle} S_x \cdot S_y$. Prove the IRB directly without appeal to RP.

As already mentioned, a successful attempt in this direction has been made by Sakai [90], who managed to apply the lace expansion to a modified random current representation of the Ising model. However, here we have in mind something perhaps more robust which addresses directly the principal reason why we need RP, which is that

the spins (S_x) are not *a priori* independent Gaussian

Among approaches in this direction is the *spherical approximation* for the $O(n)$ model, in which the constraint $|S_x| = 1$ at every spin is replaced by a constraint on $\sum_x |S_x|^2$.

The IRB is often viewed as a rigorous version of *spin-wave theory*. This theory, initiated in the work of Dyson [37] and others, describes continuous deformations of the lowest energy states by means of an appropriate Gaussian field theory. In Chapter 6 we saw that chessboard estimates may be applied *in conjunction* with spin-wave calculations — which are generally deemed to be the realm of the IRB — to prove phase

transitions. This was possible because spin-waves disqualified all but a finite number of ground states from candidacy for low-temperature states. Notwithstanding, one might be able to do the same even in the presence of infinitely many low-temperature states:

Problem 8.2. Prove symmetry breaking at low temperatures in systems with continuous internal symmetry — e.g., the $O(2)$-model — without the use of the IRB. Chessboard estimates are allowed.

An interesting resource for thinking about this problem may be the paper of Bricmont and Fontaine [21].

Further motivation to look at this problem comes from quantum theory: The quantum Heisenberg ferromagnet is not RP (see Speer [99]) and so there is no proof of the IRB and, consequently, no proof of low-temperature symmetry breaking. On the other hand, the classical Heisenberg ferromagnet is RP and so the spin-condensation argument applies. However, if we had a more robust proof of symmetry breaking in the classical model, e.g., using chessboard estimates, one might hope to extend the techniques of Biskup, Chayes and Starr [15] to include also the quantum system.

While the theory described in these notes is not restricted exclusively to ferromagnetic systems, in order to have the IRB one needs a good deal of attractivity in the system. It is actually clear that the IRB cannot hold as stated for antiferromagnets, e.g., hard core lattice gas, which is a model with variables $n_x \in \{0, 1\}$ and the "Gibbs" weight proportional to

$$\lambda^{\sum_x n_x} \prod_{\langle x,y \rangle} (1 - n_x n_y), \tag{8.1}$$

or the q-state Potts antiferromagnet, which is the model in (2.7) with $J < 0$. Indeed, the *staggered* long-range order, which is known to occur in the hard core lattice gas once $\lambda \gg 1$, implies that the macroscopically occupied mode is $k = (\pi, \dots, \pi)$ rather than $k = 0$. Nevertheless, we hope that some progress can be made and so we pose:

Problem 8.3. Derive a version of the IRB for the hard-core lattice gas and/or the q-state Potts antiferromagnet at zero temperature.

Solving this problem would, hopefully, also provide an easier passage to the proof that the critical λ for the appearance of staggered order tends to zero as $d \to \infty$ — in fact, if the mean-field theory is right then one should have $\lambda_c \sim c/d$ — and that the 3-coloring of $\mathbb{Z}^2$ exhibits six distinct extremal measures of maximal entropy. These results have recently been obtained by sophisticated contour-counting arguments [54, 55].

References

1. D.B Abraham and O.J. Heilmann, *Interacting dimers on the simple cubic lattice as a model for liquid crystals*, J. Phys. A: Math. Gen. **13** (1980) 1051–1062.
2. M. Aizenman, *Geometric analysis of ϕ^4 fields and Ising models. I, II.*, Commun. Math. Phys. **86** (1982) 1–48.
3. M. Aizenman, J.T. Chayes, L. Chayes and C.M. Newman, *Discontinuity of the magnetization in one-dimensional $1/|x-y|^2$ Ising and Potts models*, J. Statist. Phys. **50** (1988), no. 1-2, 1–40.
4. M. Aizenman and R. Fernández, *On the critical behavior of the magnetization in high-dimensional Ising models*, J. Statist. Phys. **44** (1986) 393–454.
5. M. Aizenman and R. Fernández, *Critical exponents for long-range interactions* Lett. Math. Phys. **16** (1988), no. 1, 39–49.
6. N. Angelescu and V.A. Zagrebnov, *A lattice model of liquid crystals with matrix order parameter*, J. Phys. A **15** (1982), no. 11, L639–L643.
7. M. Biskup, *Reflection positivity of the random-cluster measure invalidated for non-integer q*, J. Statist. Phys. **92** (1998) 369–375.
8. M. Biskup, C. Borgs, J.T. Chayes and R. Kotecký, *Gibbs states of graphical representations of the Potts model with external fields*, J. Math. Phys. **41** (2000), no. 3, 1170–1210.
9. M. Biskup and L. Chayes, *Rigorous analysis of discontinuous phase transitions via mean-field bounds*, Commun. Math. Phys. **238** (2003), no. 1-2, 53–93.
10. M. Biskup, L. Chayes and N. Crawford, *Mean-field driven first-order phase transitions in systems with long-range interactions*, J. Statist. Phys. **122** (2006), no. 6, 1139–1193.
11. M. Biskup, L. Chayes, and S.A. Kivelson, *Order by disorder, without order, in a two-dimensional spin system with $O(2)$-symmetry*, Ann. Henri Poincaré **5** (2004), no. 6, 1181–1205.
12. M. Biskup, L. Chayes, and R. Kotecký, *Coexistence of partially disordered/ordered phases in an extended Potts model*, J. Statist. Phys. **99** (2000), no. 5/6, 1169–1206.
13. M. Biskup, L. Chayes, and Z. Nussinov, *Orbital ordering in transition-metal compounds: I. The 120-degree model*, Commun. Math. Phys. **255** (2005), no. 2, 253–292.
14. M. Biskup, L. Chayes and Z. Nussinov, *Orbital ordering in transition-metal compounds: II. The orbital compass model*, in preparation.
15. M. Biskup, L. Chayes and S. Starr, *Quantum spin systems at positive temperature*, Commun. Math. Phys. **269** (2007), no. 3, 611–657
16. M. Biskup and R. Kotecký, *Forbidden gap argument for phase transitions proved by means of chessboard estimates*, Commun. Math. Phys. **264** (2006), no. 3, 631–656.

17. M. Biskup and R. Kotecký, *Phase coexistence of gradient Gibbs states*, Probab. Theory Rel. Fields **139** (2007), no. 1-2, 1–39.
18. M. Biskup and H. Spohn, *Scaling limit for a class of gradient fields with non-convex potentials*, Ann. Probab. (to appear)
19. T. Bodineau, *Translation invariant Gibbs states for the Ising model*, Probab. Theory Related Fields **135** (2006), no. 2, 153–168.
20. C. Borgs and E. Seiler, *Quark deconfinement at high temperature: a rigorous proof*, Nucl. Phys. B **215** (1983), no. 1, 125–135.
21. J. Bricmont and J.-R. Fontaine, *Infrared bounds and the Peierls argument in two dimensions*, Commun. Math. Phys. **87** (1982/83), no. 3, 417–427.
22. J. Bricmont, H. Kesten, J.L. Lebowitz and R.H. Schonmann, *A note on the Ising model in high dimensions*, Commun. Math. Phys. **122** (1989) 597–607.
23. R.M. Burton and M. Keane, *Density and uniqueness in percolation*, Commun. Math. Phys. **121** (1989), no. 3, 501–505.
24. L. Chayes, *Mean field analysis of low dimensional systems*, Commun. Math. Phys. (to appear)
25. L. Chayes, R. Kotecký, and S. B. Shlosman. *Staggered phases in diluted systems with continuous spins*, Commun. Math. Phys. **189** (1997) 631–640.
26. L. Chayes and J. Machta, *Graphical representations and cluster algorithms. Part I: Discrete spin systems*, Physica A **239** (1997) 542–601.
27. J.G. Conlon and J.P. Solovej, *On asymptotic limits for the quantum Heisenberg model*, J. Phys. A **23** (1990), no. 14, 3199–3213.
28. P. Curie, *Propriétés magnétiques des corps a diverses températures*, Ann. de Chimie et Physique **5** (1885) 289; reprinted in *Œuvres de Pierre Curie* (Gauthier-Villars, Paris, 1908) pp. 232–334.
29. A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* (Springer Verlag, Inc., New York, 1998).
30. J.-D. Deuschel and D.W. Stroock, *Large deviations*, Pure and Applied Mathematics, 137. Academic Press, Inc., Boston, MA, 1989.
31. J. Dimock and T.R. Hurd, *Sine-Gordon revisited*, Ann. Henri Poincaré **1** (2000), no. 3, 499–541.
32. R. Dobrushin, *The description of a random field by means of conditional probabilities and conditions of its regularity*, Theor. Prob. Appl. **13** (1968) 197–224.
33. R.L. Dobrushin and S.B. Shlosman, *Phases corresponding to minima of the local energy*, Selecta Math. Soviet. **1** (1981), no. 4, 317–338.
34. R.L. Dobrushin and S. Shlosman, *Absence of breakdown of continuous symmetry in two-dimensional models of statistical physics*, Commun. Math. Phys. **42** (1975), no. 1, 31–40.
35. R.L. Dobrushin and M. Zahradník, *Phase diagrams for continuous-spin models: an extension of the Pirogov-Sinaĭ theory*, In: R.L. Dobrushin (ed.), *Mathematical problems of statistical mechanics and dynamics*, pp. 1–123, Math. Appl. (Soviet Ser.), vol. 6, Reidel, Dordrecht, 1986.

36. R. Durrett, *Probability: Theory and Examples*, Second edition. Duxbury Press, Belmont, CA, 1996.
37. F.J. Dyson, *General theory of spin-wave interactions*, Phys. Rev. **102** (1956), no. 5, 1217–1230.
38. F.J. Dyson, E.H. Lieb and B. Simon, *Phase transitions in quantum spin systems with isotropic and nonisotropic interactions*, J. Statist. Phys. **18** (1978) 335–383.
39. R.G. Edwards and A.D. Sokal, *Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm*, Phys. Rev. D **38** (1988) 2009–2012.
40. R.S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Grundlehren der Mathematischen Wissenschaften, vol. 271 (Springer-Verlag, New York, 1985).
41. A.C.D. van Enter and S.B. Shlosman, *First-order transitions for n-vector models in two and more dimensions: Rigorous proof*, Phys. Rev. Lett. **89** (2002) 285702.
42. A.C.D. van Enter and S.B. Shlosman, *Provable first-order transitions for nonlinear vector and gauge models with continuous symmetries*, Commun. Math. Phys. **255** (2005) 21–32.
43. C.M. Fortuin and P.W. Kasteleyn, *On the random cluster model. I. Introduction and relation to other models*, Physica (Amsterdam) **57** (1972) 536–564.
44. M. Freedman, L. Lovász and A. Schrijver, *Reflection positivity, rank connectivity, and homomorphism of graphs*, J. Amer. Math. Soc. **20** (2007), no. 1, 37–51.
45. J. Fröhlich, *On the triviality of $\lambda\varphi^4_d$ theories and the approach to the critical point in $d \underset{(-)}{>} 4$ dimensions*, Nucl. Phys. B **200** (1982), no. 2, 281–296.
46. J. Fröhlich, R. Israel, E.H. Lieb and B. Simon, *Phase transitions and reflection positivity. I. General theory and long-range lattice models*, Commun. Math. Phys. **62** (1978), no. 1, 1–34.
47. J. Fröhlich, R. Israel, E.H. Lieb and B. Simon, *Phase transitions and reflection positivity. II. Lattice systems with short-range and Coulomb interactions*, J. Statist. Phys. **22** (1980), no. 3, 297–347.
48. J. Fröhlich and E.H. Lieb, *Phase transitions in anisotropic lattice spin systems*, Commun. Math. Phys. **60** (1978), no. 3, 233–267.
49. J. Fröhlich and C.-E. Pfister, *On the absence of spontaneous symmetry breaking and of crystalline ordering in two-dimensional systems*, Commun. Math. Phys. **81** (1981), no. 2, 277–298.
50. J. Fröhlich, B. Simon and T. Spencer, *Infrared bounds, phase transitions and continuous symmetry breaking*, Commun. Math. Phys. **50** (1976) 79–95.
51. J. Fröhlich and T. Spencer, *The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas*, Commun. Math. Phys. **81** (1981), no. 4, 527–602.

52. T. Funaki, *Stochastic Interface Models*, Lecture Notes for the International Probability School at Saint-Flour, Lecture Notes in Math., 1869, Springer, Berlin, 2005.
53. T. Funaki and H. Spohn, *Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model*, Commun. Math. Phys. **185** (1997) 1–36.
54. D. Galvin and J. Kahn, *On phase transition in the hard-core model on $\mathbb{Z}^d$*, Combin. Probab. Comput. **13** (2004), no. 2, 137–164.
55. D. Galvin and D. Randall, *Torpid mixing of local markov chains on 3-Colorings of the discrete torus*, Proceedings of the Eighteenth Annual ACM–SIAM Symposium on Discrete Algorithms (SODA 2007), to appear.
56. P.G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, Oxford University Press, New York, 1993.
57. H.-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter Studies in Mathematics, vol. 9 (Walter de Gruyter & Co., Berlin, 1988).
58. T. Gobron and I. Merola, *First-order phase transition in Potts models with finite-range interactions*, J. Statist. Phys. **126** (2007), no. 3, 507–583.
59. R.B. Griffiths, *Peierls proof of spontaneous magnetization in a two-dimensional Ising ferromagnet*, Phys. Rev. (2) **136** (1964) A437–A439.
60. R.B. Griffiths, *Correlations in Ising ferromagnets*, J. Math. Phys. **8** (1967), no. 3, 478–483.
61. G. Grimmett, *The Random-Cluster Model*, Grundlehren der Mathematischen Wissenschaften, vol. 333, Springer-Verlag, Berlin, 2006
62. O.J. Heilmann and E.H. Lieb, *Lattice models for liquid crystals*, J. Statist. Phys. **20** (1979), no. 6, 679–693.
63. C.L. Henley, *Ordering due to disorder in a frustrated vector antiferromagnet*, Phys. Rev. Lett. **62** (1989) 2056–2059.
64. F. den Hollander, *Large deviations*, Fields Institute Monographs, 14. American Mathematical Society, Providence, RI, 2000.
65. D. Ioffe, S. Shlosman and Y. Velenik, *2D models of statistical physics with continuous symmetry: The case of singular interactions*, Commun. Math. Phys. **226** (2002), no. 2, 433–454.
66. R.B. Israel, *Convexity in the theory of lattice gases*, With an introduction by Arthur S. Wightman. Princeton University Press, Princeton, N.J., 1979.
67. T. Kennedy, *Long range order in the anisotropic quantum ferromagnetic Heisenberg model*, Commun. Math. Phys. **100** (1985), no. 3, 447–462.
68. T. Kennedy, E.H. Lieb and B.S. Shastry, *Existence of Néel order in some spin-$\frac{1}{2}$ Heisenberg antiferromagnets*, J. Statist. Phys. **53** (1988), no. 5-6, 1019–1030,
69. H. Kesten and R. Schonmann, *Behavior in large dimensions of the Potts and Heisenberg models*, Rev. Math. Phys. **1** (1990) 147–182.

70. M. Kosterlitz and D.J. Thouless, *Ordering, metastability and phase transitions in two-dimensional systems*, J. Phys. C: Solid State Phys., **6** (1973) 1181–1203.
71. R. Kotecký and S.B. Shlosman, *Existence of first-order transitions for Potts models*, In: S. Albeverio, Ph. Combe, M. Sirigue-Collins (eds.), Proc. of the International Workshop — Stochastic Processes in Quantum Theory and Statistical Physics, Lecture Notes in Physics **173**, 248–253, Springer-Verlag, Berlin-Heidelberg-New York, 1982.
72. R. Kotecký and S.B. Shlosman, *First-order phase transitions in large entropy lattice models*, Commun. Math. Phys. **83** (1982), no. 4, 493–515.
73. U. Krengel, *Ergodic theorems*, de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, 1985.
74. C. Külske, *The continuous spin random field model: Ferromagnetic ordering in $d \geq 3$*, Rev. Math. Phys. **11** (1999), no. 10, 1269–1314.
75. C. Külske, *Stability for a continuous SOS-interface model in a randomly perturbed periodic potential*, WIAS Preprint no. 466 (1998); http://www.math.rug.nl/~kuelske/publications.html
76. G.L. Lawler, *Intersections of random walks*, Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, 1991.
77. E.H. Lieb, *Two theorems on the Hubbard model*, Phys. Rev. Lett. **62** (1989), no. 10, 1201–1204. Errata: Phys. Rev. Lett. **62** (1989), no. 16, 1927.
78. E.H. Lieb, *Flux phase of the half-filled band*, Phys. Rev. Lett. **73** (1994) 2158-2161.
79. N. Macris, *Periodic ground states in simple models of itinerant fermions interacting with classical fields*, The nature of crystalline states (Kudowa-Zdrój, 1995). Phys. A **232** (1996), no. 3-4, 648–656.
80. N. Macris and J. Lebowitz, *Ground states and low-temperature phases of itinerant electrons interacting with classical fields: a review of rigorous results*, Quantum problems in condensed matter physics, J. Math. Phys. **38** (1997), no. 4, 2084–210.
81. N. Macris and B. Nachtergaele, *On the flux phase conjecture at half-filling: an improved proof*, J. Statist. Phys. 85 (1996), no. 5-6, 745–761.
82. N.D. Mermin and H. Wagner, *Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models*, Phys. Rev. Lett. **17** (1966), no. 22, 1133–1136.
83. Z. Nussinov, M. Biskup, L. Chayes and J. van den Brink, *Orbital order in classical models of transition-metal compounds*, Europhys. Lett. **67** (2004), no. 6, 990–996.
84. K. Osterwalder and R. Schrader, *Axioms for Euclidean Green's functions*, Commun. Math. Phys. **31** (1973), 83–112.
85. A. Patrasciou and E. Seiler, *Percolation theory and the existence of a soft phase in 2D spin models*, Nucl. Phys. B (Proc. Suppl.) **30** (1993) 184–191.

86. K. Petersen, *Ergodic theory*, Cambridge Studies in Advanced Mathematics, 2. Cambridge University Press, Cambridge, 1989.
87. C.-E. Pfister, *On the symmetry of the Gibbs states in two-dimensional lattice systems*, Commun. Math. Phys. **79** (1981), no. 2, 181–188.
88. D. Ruelle, *Thermodynamic formalism. The mathematical structures of equilibrium statistical mechanics*, Second edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 200
89. D. Ruelle, *Statistical mechanics. Rigorous results*, Reprint of the 1989 edition, World Scientific Publishing Co., Inc., River Edge, NJ; Imperial College Press, London, 1999.
90. A. Sakai, *Lace expansion for the Ising model*, Commun. Math. Phys. **272** (2007), no. 2, 283–344.
91. M. Salmhofer and E. Seiler, *Proof of chiral symmetry breaking in strongly coupled lattice gauge theory*, Commun. Math. Phys. **139** (1991), no. 2, 395–432. Errata: *ibid* **146** (1992), no. 3, 637–638.
92. S. Sheffield, *Random Surfaces*, Asterisque 2005, No. 304, 177pp.
93. E.F. Shender, *Antiferromagnetic garnets with fluctuationally interacting sublattices*, Sov. Phys. JETP **56** (1982) 178–184.
94. S. Shlosman, *Phase transitions for two-dimensional models with isotropic short range interactions and continuous symmetry*, Commun. Math. Phys. **71** (1980) 207–212.
95. S.B. Shlosman, *The method of reflective positivity in the mathematical theory of phase transitions of the first kind* (Russian), Uspekhi Mat. Nauk **41** (1986), no. 3(249), 69–111, 240.
96. S. Shlosman and Y. Vignaud, *Dobrushin interfaces via reflection positivity*, Commun. Math. Phys. 276 (2007), no. 3, 827–86.
97. B. Simon, *The statistical mechanics of lattice gases*, Vol. I., Princeton Series in Physics (Princeton University Press, Princeton, NJ, 1993).
98. A.D. Sokal, *An alternate constructive approach to the φ_3^4 quantum field theory, and a possible destructive approach to φ_4^4*, Ann. Inst. H. Poincaré Sect. A **37** (1982), no. 4, 317–398.
99. E.R. Speer, *Failure of reflection positivity in the quantum Heisenberg ferromagnet*, Lett. Math. Phys. **10** (1985), no. 1, 41–47.
100. F. Spitzer, *Principles of random walks*, Second edition. Graduate Texts in Mathematics, Vol. 34. Springer-Verlag, New York-Heidelberg, 1976.
101. H. Tasaki, *Ferromagnetism in the Hubbard model: a constructive approach*, Commun. Math. Phys. **242** (2003), no. 3, 445–472.
102. G.-S. Tian, *Lieb's spin-reflection-positivity method and its applications to strongly correlated electron systems*, J. Statist. Phys. **116** (2004), no. 1-4, 629–680.
103. B. Tóth's website:
`www.math.bme.hu/~balint/oktatas/statisztikus_fizika/jegyzet/`
104. Y. Velenik, *Localization and delocalization of random interfaces*, Probab. Surveys **3** (2006) 112–169.

105. P. Weiss, *L'hypothèse du champ moléculaire et la propriété ferromagnétique*, J. de Physique **6** (1907) 661–689.
106. F.Y. Wu, *The Potts model*, Rev. Modern Phys. **54** (1982) 235–268.
107. M. Zahradník, *Contour methods and Pirogov-Sinai theory for continuous spin lattice models*, In: R.A. Minlos, S. Shlosman and Yu.M. Suhov (eds.), *On Dobrushin's way. From probability theory to statistical physics*, pp. 197–220, Amer. Math. Soc. Transl. Ser. 2, vol. 198, Amer. Math. Soc., Providence, RI, 2000.

Stochastic Geometry of Classical and Quantum Ising Models

Dmitry Ioffe

Faculty of Industrial Engineering and Management, Technion, Haifa, Israel
ieioffe@ie.technion.ac.il

Contents

1 Introduction

These lecture notes are based on a mini-course which I taught at Prague school in September 2006. The idea was to try to develop and explain to probabilistically minded students a unified approach to the

R. Kotecký (ed.), *Methods of Contemporary Mathematical Statistical Physics*, Lecture Notes in Mathematics 1970, DOI: 10.1007/978-3-540-92796-9_2,

Fortuin-Kasteleyn (FK) and to the random current (RC) representation of classical and quantum Ising models via path integrals. No background in quantum statistical mechanics was assumed.

In Section 1 familiar classical Ising models are rewritten in the quantum language. In this way usual FK and RC representations emerge as different instances of Lie-Trotter product formula. Then I am following [4] and set up a general notation for the Poisson limits.

In Section 2 both FK and the RC representations are generalized to quantum Ising models in transverse field. The FK representation was originally derived in [8] and [3]. The observation regarding the RC representation seems to be new. Both representations are used to derive formulas for one and two point functions and for the matrix and reduced density matrix elements.

Section 3 is devoted to the quantum Curie-Weiss model in transverse field. In the quantum mean field case the FK representation is built upon a generalization of the classical random graph model. I briefly explain recent results of [15], where the critical curve for quantum random graphs was explicitly computed. The critical curve for the quantum Curie-Weiss model itself is computed in the concluding Subsection 3.3 via partial Trotterization and a large deviation approach.

Of course, stochastic geometric methods apply for a large variety of other models, see the seminal [4] as well as [18, 20] and references therein. I did not try to provide a complete bibliography on the subject - the emphasis was rather on trying to advertise probabilistic aspects of quantum spin systems to a reader who is (like me) not very well familiar with the latter. I, therefore, apologize for many excellent and relevant papers which were not mentioned.

2 Classical Ising Model

We use the following notation for the classical Ising model:

- $(\Lambda, \mathcal{E})$ is a finite graph with *unoriented* edges $e = \{i, j\} = \{j, i\} \in \mathcal{E}$.
- $\mathbf{J} = \{J_{ij} \geq 0\}$ are coupling constants. By definition $J_{ij} > 0 \Leftrightarrow \{i, j\} \in \mathcal{E}$.
- $h \in \mathbb{R}$ is a magnetic field.
- $\nu \in \Omega_\Lambda \stackrel{\Delta}{=} \{-1, 1\}^\Lambda$ is a spin configuration on Λ.

The Hamiltonian $\mathbf{H}_\Lambda$ is a function on Ω_Λ,

$$-\mathbf{H}_\Lambda(\nu) = \sum_{(i,j)\in\mathcal{E}} J_{ij}\nu_i\nu_j + h\sum_{i\in\Lambda}\nu_i.$$

Given $\beta \geq 0$ (inverse temperature) define the classical Ising-Gibbs probability distribution $\mu_{\Lambda}^{\beta,h}$ on Ω_{Λ} as

$$\mu_{\Lambda}^{\beta,h}(\nu) \;=\; \frac{1}{\mathcal{Z}_{\Lambda}(\beta,h)} \mathrm{e}^{-\beta \mathbf{H}_{\Lambda}(\nu)},$$

where the normalizing constant (partition function) is given by

$$\mathcal{Z}_{\Lambda}(\beta,h) \;=\; \sum_{\nu \in \Omega_{\Lambda}} \mathrm{e}^{-\beta \mathbf{H}_{\Lambda}(\nu)}. \tag{2.1}$$

In the sequel we shall use $\mu_{\Lambda}^{\beta,h}(\bullet)$ for the expectation under $\mu_{\Lambda}^{\beta,h}$. In particular, the mean value of the spin at i is

$$\mu_{\Lambda}^{\beta,h}(\nu_i) \;=\; \frac{1}{\mathcal{Z}_{\Lambda}(\beta,h)} \sum_{\nu \in \Omega_{\Lambda}} \nu_i \mathrm{e}^{-\beta \mathbf{H}_{\Lambda}(\nu)}, \tag{2.2}$$

and the two-point function $\mu_{\Lambda}^{\beta,h}(\nu_i \nu_j)$ is

$$\mu_{\Lambda}^{\beta,h}(\nu_i \nu_j) \;=\; \frac{1}{\mathcal{Z}_{\Lambda}(\beta,h)} \sum_{\nu \in \Omega_{\Lambda}} \nu_i \nu_j \mathrm{e}^{-\beta \mathbf{H}_{\Lambda}(\nu)}. \tag{2.3}$$

Two examples we shall consider in this paper are:

1. Curie-Weiss model: $\Lambda = \{1, 2 \ldots, N\}$ and $J_{ij} \equiv 1/N$.
2. Finite range Ising model: $\Lambda \subset \mathbb{Z}^d$ and $J_{ij} = 0$ for $\|i - j\| \geq R$.

2.1 Classical Ising Model Dressed as Quantum

Let us re-derive formulas (2.1), (2.2) and (2.3) in the quantum language. In this way spin values ± 1 are understood as eigenvalues of Pauli matrix

$$\hat{\sigma}^{\mathsf{z}} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.4}$$

Let us define the corresponding eigenfunctions

$$\psi_{+1} \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \psi_{-1} \;=\; \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.5}$$

Of course, $\hat{\sigma}^{\mathsf{z}} \psi_{\nu} = \nu \psi_{\nu}$ for $\nu = \pm 1$. Throughout these lectures we shall work only with real numbers. Using eigenfunctions $\psi_{\pm 1}$ one constructs the following "lifting" of classical configurations $\nu \in \Omega_{\Lambda}$: Define

$$\mathbb{X}_{\Lambda} \;=\; \bigotimes_{i \in \Lambda} \mathbb{R}^2.$$

$\mathbb{X}_\Lambda$ is a $2^{|\Lambda|}$-dimensional vector space over the field of reals. Classical configurations $\nu \in \Omega_\Lambda$ are encoded in $\mathbb{X}_\Lambda$ as tensor products,

$$\Psi_\nu \stackrel{\Delta}{=} \otimes_{i\in\Lambda}\psi_{\nu_i}. \tag{2.6}$$

The collection $\{\Psi_\nu\}_{\nu\in\Omega_\Lambda}$ is a complete orthonormal basis of $\mathbb{X}_\Lambda$ with respect to the scalar product

$$\langle\Psi_\nu|\Psi_{\nu'}\rangle \stackrel{\Delta}{=} \prod_{i\in\Lambda}\langle\psi_{\nu_i},\psi_{\nu'_i}\rangle_2,$$

where $\langle\bullet,\bullet\rangle_2$ is the usual scalar product of $\mathbb{R}^2$. With each $i \in \Lambda$ we associate a linear self-adjoint operator (symmetric matrix) $\hat\sigma^{\mathsf{z}}_i$ which acts on i-th coordinate of Ψ as a copy of Pauli matrix $\hat\sigma^{\mathsf{z}}$ defined in (2.4). Namely, for each $\nu \in \Omega_\Lambda$,

$$\hat\sigma^{\mathsf{z}}_i\Psi_\nu \stackrel{\Delta}{=} \psi_{\nu_1}\otimes\cdots\otimes\hat\sigma^{\mathsf{z}}\psi_{\nu_i}\otimes\ldots = \nu_i\Psi_\nu. \tag{2.7}$$

Obviously, $\hat\sigma^{\mathsf{z}}_i$ and $\hat\sigma^{\mathsf{z}}_j$ commute, and, moreover,

$$\hat\sigma^{\mathsf{z}}_i\hat\sigma^{\mathsf{z}}_j\Psi_\nu = \nu_i\nu_j\Psi_\nu. \tag{2.8}$$

Define now the quantum Hamiltonian $\mathcal{H}_\Lambda$ as a linear self-adjoint operator on $\mathbb{X}_\Lambda$,

$$-\mathcal{H}_\Lambda = \sum_{(i,j)\in\mathcal{E}} J_{ij}\hat\sigma^{\mathsf{z}}_i\hat\sigma^{\mathsf{z}}_j + h\sum_{i\in\Lambda}\hat\sigma^{\mathsf{z}}_i. \tag{2.9}$$

Then, (2.7) and (2.8) imply,

$$\mathcal{H}_\Lambda\Psi\nu = \mathbf{H}_\Lambda(\nu)\Psi_\nu.$$

In other words, $\mathcal{H}_\Lambda$ is diagonal in the $\{\Psi_\nu\}$ basis, and with the corresponding eigenvalues being equal to values of the classical Ising Hamiltonian on configurations ν.

It is possible now to rewrite classical formulas (2.1)-(2.3) in terms of the quantum Hamiltonian $\mathcal{H}_\Lambda$. First of all,

$$\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right) = \sum_{\nu\in\Omega_\Lambda}\langle\Psi_\nu|\mathrm{e}^{-\beta\mathcal{H}_\Lambda}|\Psi_\nu\rangle = \sum_{\nu\in\Omega_\Lambda}\mathrm{e}^{-\beta\mathbf{H}_\Lambda(\nu)} = \mathcal{Z}_\Lambda(\beta,h). \tag{2.10}$$

Similarly,

$$\mu^{\beta,h}_\Lambda(\nu_i) = \frac{\mathsf{Tr}\left(\hat\sigma^{\mathsf{z}}_i\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)}{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)} \quad\text{and}\quad \mu^{\beta,h}_\Lambda(\nu_i\nu_j) = \frac{\mathsf{Tr}\left(\hat\sigma^{\mathsf{z}}_i\hat\sigma^{\mathsf{z}}_j\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)}{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)}. \tag{2.11}$$

2.2 Path Integral Representation and Poisson Limits

Since all the operators $\{\hat{\sigma}_i^z\}$ commute,

$$\mathrm{e}^{-\beta\mathcal{H}_\Lambda} \;=\; \left(\prod_{(i,j)} \mathrm{e}^{\Delta J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z} \prod_i \mathrm{e}^{\Delta h\hat{\sigma}_i^z}\right)^{\beta/\Delta}. \tag{2.12}$$

To facilitate the exposition we shall focus now on the case of zero magnetic field $h = 0$, the full Hamiltonian with both non-zero h will be considered in Subsection 2.3 and in Subsection 2.4, furthermore, an additional positive field in the traverse direction will be considered in Section 3.

For small Δ we shall linearize $\mathrm{e}^{\Delta J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z}$ in (2.12) in two different ways:

1) Write

$$J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z \;=\; J_{ij}\mathrm{I} - J_{ij}\mathrm{I} + J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z.$$

Then,

$$\mathrm{e}^{-\beta\mathcal{H}_\Lambda} \;=\; \mathrm{e}^{\beta\sum_{(i,j)} J_{ij}} \lim_{\Delta\to 0}\left(\prod_{(i,j)} \left\{(1-\Delta J_{ij})\mathrm{I} + \Delta J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z\right\}\right)^{\beta/\Delta}. \tag{2.13}$$

This will lead to the random current representation of the model.

2) Write

$$J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z \;=\; J_{ij}\mathrm{I} - 2J_{ij}\mathrm{I} + 2J_{ij}\frac{\mathrm{I}+\hat{\sigma}_i^z\hat{\sigma}_j^z}{2}.$$

In the latter case,

$$\mathrm{e}^{-\beta\mathcal{H}_\Lambda} \;=\; \mathrm{e}^{\beta\sum_{(i,j)} J_{ij}} \lim_{\Delta\to 0}\left(\prod_{(i,j)} \left\{(1-2\Delta J_{ij})\mathrm{I} + 2\Delta J_{ij}\frac{\mathrm{I}+\hat{\sigma}_i^z\hat{\sigma}_j^z}{2}\right\}\right)^{\beta/\Delta}. \tag{2.14}$$

As we shall see below such linearization leads to the FK (Fortuin-Kasteleyn) representation of the model. Thus both the FK and the random current representations are instances of path integral representation via Poisson limits which, following [4], we proceed to discuss in a somewhat general context.

General Setup for Poisson Limits

The fact that the operators $\hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}}$ in (2.13) or operators $(\mathrm{I}+\hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}})/2$ in (2.14) commute *is not* essential for the path integral representation via Poisson limits. For the rest of this Subsection we shall work in the following general context:

1. $\mathbb{X}$ is an M-dimensional vector space (over $\mathbb{R}$) with a scalar product $\langle\bullet|\bullet\rangle$ and an orthonormal basis $\{\Psi_i\}$
2. $K_1,\dots,K_m$ are self-adjoint operators (matrices) on $\mathbb{X}$, in general non-commuting.
3. $\lambda_1,\dots,\lambda_m$ are positive numbers.

Given $\beta>0$, we would like to find a probabilistic representation for

$$\exp\Big\{\beta\sum_1^m\lambda_l K_l\Big\} \tag{2.15}$$

The linearization relies on two basic facts from theory of matrices:

Lie-Trotter Formula

Let A and B be two matrices. Then

$$\mathrm{e}^{A+B} = \lim_{n\to\infty}\Big(\mathrm{e}^{A/n}\mathrm{e}^{B/n}\Big)^n. \tag{2.16}$$

Proof (following [19]). Set

$$T_n=\mathrm{e}^{(A+B)/n} \qquad \text{and} \qquad S_n=\mathrm{e}^{A/n}\mathrm{e}^{B/n}.$$

Then,

$$T_n^n-S_n^n=\sum_{l=0}^{n-1}\Big(T_n^{n-l}S_n^l-T_n^{n-l-1}S_n^{l+1}\Big)=\sum_{l=0}^{n-1}T_n^{n-l-1}\,(T_n-S_n)\,S_n^l.$$

Now,

$$T_n-S_n=\sum_{k=0}^{\infty}\frac{1}{k!}\Big(\frac{A+B}{n}\Big)^k-\Big\{\sum_{k=0}^{\infty}\frac{1}{k!}\Big(\frac{A}{n}\Big)^k\Big\}\Big\{\sum_{k=0}^{\infty}\frac{1}{k!}\Big(\frac{B}{n}\Big)^k\Big\}=O(\frac{1}{n^2}).$$

On the other hand, $\|T_n^{n-l-1}\|\cdot\|S_n^l\|\le\mathrm{e}^{\|A\|+\|B\|}$ for all $l=1,\dots,n-1$. □

Product Expansion Formula

Let $A_1, A_2, \ldots, A_n$ be self-adjoint matrices and Ψ, Ψ' two vectors in $\mathbb{X}$. Then,

$$\langle \Psi | A_1 \ldots A_n | \Psi' \rangle \;=\; \sum_{\Psi_{i_1}, \ldots \Psi_{i_n}} \langle \Psi | A_1 | \Psi_{i_1} \rangle \langle \Psi_{i_1} | A_2 | \Psi_{i_2} \rangle \ldots \langle \Psi_{i_{n-1}} | A_n | \Psi' \rangle, \tag{2.17}$$

where Ψ_{i_l}-s run through the elements of the orthonormal basis $\{\Psi_i\}$ for all $l = 1, \ldots, n-1$.

Proof. In the case $n = 2$ (2.17) follows from expansion of $A_1\Psi$ in the basis $\{\Psi_i\}$,

$$A_1\Psi \;=\; \sum_{i_1=1}^{M} \langle \Psi | A_1 | \Psi_{i_1} \rangle \Psi_{i_1}.$$

The general case follows by induction. □

Path Integral Representation

Let us go back to (2.15). By Lie-Trotter formula (2.16),

$$\exp\left\{ \beta \sum_1^m \lambda_l K_l \right\} \;=\; \mathrm{e}^{\beta \sum \lambda_l} \lim_{\Delta \to 0} \left(\prod_{l=1}^{m} \{(1 - \Delta\lambda_l)\mathrm{I} + \Delta\lambda_l K_l\} \right)^{\beta/\Delta}. \tag{2.18}$$

In the sequel we shall tacitly assume that $\beta/\Delta \in \mathbb{N}$. For each $l = 1, \ldots, m$ consider a sequence of iid Bernoulli random variables

$$\underline{\xi}_l = \{\xi_l(1), \xi_l(2) \ \ldots, \xi_l(\beta/\Delta)\},$$

with the probability of success being equal to $\Delta\lambda_l$. We assume that the sequences $\underline{\xi}_l$ are independent and let $\mathbb{P}^{\underline{\lambda}}_{\beta,\Delta}$ be the corresponding probability measure on

$$\{0,1\}^{\beta/\Delta} \times \cdots \times \{0,1\}^{\beta/\Delta}$$

Above $\underline{\lambda}$ is a shorthand notation for the vector of success rates $\{\lambda_1, \ldots, \lambda_m\}$. Then we can expand the expression on the right-hand side of (2.18) as follows,

$$\left(\prod_{l=1}^{m} \{(1 - \Delta\lambda_l)\mathrm{I} + \Delta\lambda_l K_l\} \right)^{\beta/\Delta} = \sum_{\underline{a}_1, \ldots \underline{a}_m} \mathbb{P}^{\underline{\lambda}}_{\beta,\Delta} \left(\bigcap_{l=1}^{m} \left\{ \underline{\xi}_l = \underline{a}_l \right\} \right) \mathcal{K}_{\underline{a}}, \tag{2.19}$$

where the matrix $\mathcal{K}_{\underline{a}}$ is defined by

$$\mathcal{K}_{\underline{a}} \stackrel{\Delta}{=} \mathcal{K}_{\underline{a}_1,\ldots,\underline{a}_m} = \prod_{j=1}^{\beta/\Delta}\Big\{\prod_{l=1}^{m}((1-a_l(j))\mathrm{I} + a_l(j)K_l)\Big\}. \tag{2.20}$$

Our next step is to associate with each sequence $\underline{\xi}_l$ of Bernoulli trials a point process of arrivals of operators K_l on the interval $[0,\beta]$. Define,

$$\xi_l^{\Delta} = \sum_{j=1}^{\beta/\delta}\xi_l(j)\delta_{i\Delta}. \tag{2.21}$$

Let Ψ and Ψ' be two elements of the basis $\{\Psi_i\}$. In order to derive a path integral representation of $\langle\Psi|K_{\underline{a}_1},\ldots,K_{\underline{a}_m}|\Psi'\rangle$ notice first of all that up to probabilities of order $O(\Delta)$ we may restrict attention to sequences $\underline{a}_1,\ldots,\underline{a}_m$ with disjoint occurrence of successes, that is $\sum_l a_l(j) = 0$ or 1 for every $j = 1,\ldots,\beta/\Delta$. In the language of (2.21) this means that the realizations of $\xi_1^{\Delta},\ldots,\xi_m^{\Delta}$ are pairwise disjoint and hence for each arrival time

$$t \in \xi^{\Delta} \stackrel{\Delta}{=} \cup\xi_l^{\Delta} = \left\{j\Delta : \sum_{l=1}^{m}a_l(j) = 1\right\},$$

there is a well defined arrival type $\mathfrak{l}^{\Delta}(t) \in \{1,\ldots,m\}$. Accordingly, one can rewrite

$$\mathcal{K}_{\underline{a}_1,\ldots,\underline{a}_m} = \prod_{j=1}^{\beta/\Delta}\left\{\delta_{\{j\Delta\notin\xi^{\Delta}\}}\mathrm{I} + \delta_{\{j\Delta\in\xi^{\Delta}\}}K_{\mathfrak{l}^{\Delta}(j\Delta)}\right\} \stackrel{\Delta}{=} \prod_{j=1}^{\beta/\Delta}\tilde{K}_j^{\Delta}.$$

By the product expansion formula (2.17),

$$\langle\Psi|\mathcal{K}_{\underline{a}}|\Psi'\rangle = \sum_{\Psi_{i_1},\ldots\Psi_{i_{\beta/\Delta-1}}}\langle\Psi|\tilde{K}_1^{\Delta}|\Psi_{i_1}\rangle\prod_{j=2}^{\beta/\Delta-1}\langle\Psi_{i_{j-1}}|\tilde{K}_j^{\Delta}|\Psi_{i_j}\rangle\langle\Psi_{i_{\beta/\Delta-1}}|\tilde{K}_{\beta/\Delta}^{\Delta}|\Psi'\rangle. \tag{2.22}$$

Of course,

$$\langle\Psi_l|\tilde{K}_j^{\Delta}|\Psi_k\rangle = \begin{cases}\delta_{\{\Psi_l=\Psi_k\}} & \text{if } j\Delta\notin\xi^{\Delta}\\ \langle\Psi_l|K_{\mathfrak{l}^{\Delta}(j\Delta)}|\Psi_k\rangle & \text{if } j\Delta\in\xi^{\Delta}\end{cases} \tag{2.23}$$

We can now put this into the continuous time context as follows: To a given sequence $\Psi,\Psi_{i_1},\ldots,\Psi_{i_{\beta/\Delta-1}},\Psi'$ associate a piecewise constant

function $\Psi^\Delta : [0,\beta] \to \{\Psi_j\}$, such that $\Psi^\Delta = \Psi$ on $[0,\Delta)$, $\Psi^\Delta(\beta) = \Psi'$, and,

$$\Psi^\Delta \,=\, \Psi_{i_j} \quad \text{on} \quad [j\Delta,(j+1)\Delta) \quad \text{for } j = 1,\dots\beta/\Delta - 1.$$

Given a realization ξ^Δ let us say that a piecewise constant function Ψ^Δ as above is compatible with ξ^Δ, $\Psi^\Delta \sim \xi^\Delta$ if all the jumps of Ψ^Δ occur only at arrival times of ξ^Δ. By (2.23) only compatible functions contribute to (2.22). In fact, in the notation just introduced the latter expansion reads as,

$$\langle\Psi|\mathcal{K}_{\underline{a}_1,\dots,\underline{a}_m}|\Psi'\rangle \,=\sum_{\Psi^\Delta\sim\xi^\Delta}\prod_{t\in\xi^\Delta}\langle\Psi^\Delta(t-)|K_{\mathfrak{l}^\Delta(t)}|\Psi^\Delta(t)\rangle. \tag{2.24}$$

Poisson Limits

A basic result on Poisson approximation implies that

$$\left(\xi_1^\Delta,\dots,\xi_m^\Delta,\mathfrak{l}^\Delta\right) \,\Rightarrow\, (\xi_1,\dots,\xi_m,\mathfrak{l})$$

where $(\xi_1,\dots,\xi_m)$ are independent Poisson point processes on $[0,\beta]$ with intensities $(\lambda_1,\dots,\lambda_m)$ respectively. Let us use $\mathbb{P}_\beta^{\underline{\lambda}}$ for the distribution of the latter. By independence there are no simultaneous arrivals, that is the type $\mathfrak{l}(t) \in \{1,\dots,m\}$ of an arrival is well defined for each $t \in \xi \stackrel{\Delta}{=} \cup\xi_l$. Furthermore, conditioned on the realization of ξ the arrival types $\mathfrak{l}(t)$ are independent and

$$\mathbb{P}_\beta^{\underline{\lambda}}\left(\mathfrak{l}(t) = l \,\middle|\, t\in\xi\right) \,=\, \frac{\lambda_l}{\lambda_1+\dots+\lambda_m}.$$

Passing to the limit $\Delta \to 0$ in (2.24) and (2.19), we arrive to the following representation of matrix elements of $\exp\{\beta\sum\lambda_l K_l\}$: For every two elements of the basis $\Psi,\Psi' \in \{\Psi_i\}$,

$$\frac{\langle\Psi|\mathrm{e}^{\beta\sum\lambda_l K_l}|\Psi'\rangle}{\exp\{\beta\sum_l\lambda_l\}} \,=\, \int\mathbb{P}_\beta^{\underline{\lambda}}(\mathrm{d}\xi_1\dots\mathrm{d}\xi_m)\sum_{\Psi\sim\xi}\prod_{t\in\xi}\langle\Psi(t-)|K_{\mathfrak{l}(t)}|\Psi(t)\rangle, \tag{2.25}$$

where, given a realization of ξ the summation is over all ξ-compatible (having jumps only at arrival times of ξ) piecewise constant right-continuous functions $\Psi : [0,\beta] \mapsto \{\Psi_i\}$, which, in addition, satisfy boundary conditions $\Psi(0) = \Psi$ and $\Psi(\beta) = \Psi'$. Clearly, since $\mathbb{X}$ is finite dimensional, and since, there are $\mathbb{P}_\beta^{\underline{\lambda}}$-a.s. finite number of arrivals of ξ, there are $\mathbb{P}_\beta^{\underline{\lambda}}$-a.s. finitely many such compatible functions.

Formula (2.25) enables a re-interpretation of various quantities related to the Hamiltonians $\mathcal{H}$ in terms of stochastic geometry of the family of Poisson processes ξ. For example,

$$\frac{\mathsf{Tr}\left(\mathrm{e}^{\beta\sum\lambda_l K_l}\right)}{\exp\left\{\beta\sum_l\lambda_l\right\}} \;=\; \int\mathbb{P}^{\lambda}_{\beta}(\mathrm{d}\xi)\sum_{\Psi\sim\xi}\langle\Psi(0)|\Psi(\beta)\rangle\prod_{t\in\xi}\langle\Psi(t-)|K_{\mathrm{l}(t)}|\Psi(t)\rangle. \tag{2.26}$$

In general, given a self-adjoint matrix A,

$$\frac{\mathsf{Tr}\left(A\mathrm{e}^{\beta\sum\lambda_l K_l}\right)}{\exp\left\{\beta\sum_l\lambda_l\right\}} \;=\; \int\mathbb{P}^{\lambda}_{\beta}(\mathrm{d}\xi)\sum_{\Psi\sim\xi}\langle\Psi(0)|A|\Psi(\beta)\rangle\prod_{t\in\xi}\langle\Psi(t-)|K_{\mathrm{l}(t)}|\Psi(t)\rangle. \tag{2.27}$$

The approach has two degrees of freedom to play with:

1. There are different ways to decompose $\mathcal{H}$ as $\mathcal{H}=-\sum\lambda_l K_l$.
2. There are different choices of orthonormal bases $\{\Psi_i\}$ of $\mathbb{X}$.

In the following two subsections we shall consider the FK and the RC (random current) representation of classical Ising systems (2.9) as different instances of the path integral representation (2.25). Then in Section 3 we shall develop the FK and the RC representation for genuine quantum systems in traverse magnetic field.

2.3 FK Representation

Classical FK representation corresponds to the decomposition of the Hamiltonian $\mathcal{H}_\Lambda$ in (2.9) as,

$$-\mathcal{H}_\Lambda \;=\; -\left(\sum_{(i,j)}J_{ij}+\sum_i h\right)\mathrm{I}+\sum_{(i,j)}2J_{ij}\frac{\mathrm{I}+\hat\sigma^{\mathsf{z}}_i\hat\sigma^{\mathsf{z}}_j}{2}+\sum_i 2h\frac{\mathrm{I}+\hat\sigma^{\mathsf{z}}_i}{2},$$

with matrix elements of $\mathrm{e}^{-\beta\mathcal{H}_\Lambda}$ being computed in the z-basis (2.6).

In the language of the preceeding Subsection, we are dealing with independent Poisson processes ξ_{ij} of arrivals of operators $K_{ij}\stackrel{\Delta}{=}\frac{\mathrm{I}+\hat\sigma^{\mathsf{z}}_i\hat\sigma^{\mathsf{z}}_j}{2}$ with intensities $2J_{ij}$ and with independent Poisson processes ξ_i of arrivals of operators $K_i\stackrel{\Delta}{=}\frac{\mathrm{I}+\hat\sigma^{\mathsf{z}}_i}{2}$ with intensities $2h$ each. Let $\nu,\nu'\in\Omega_\Lambda$ be two classical configurations and let, as before, Ψ_ν and $\Psi_{\nu'}$ be the corresponding elements of the basis of $\mathbb{X}_\Lambda$. Then,

$$\langle\Psi_\nu|K_{ij}|\Psi_{\nu'}\rangle \;=\; \delta_{\{\nu=\nu'\}}\delta_{\{\nu_i=\nu_j\}}. \tag{2.28}$$

Similarly,

$$\langle \Psi_\nu | K_i | \Psi_{\nu'} \rangle \;=\; \delta_{\{\nu=\nu'\}} \delta_{\{\nu_i=1\}}. \tag{2.29}$$

Due to our choice of the orthonormal basis, any piecewise constant function $\Psi : [0,\beta] \mapsto \{\Psi_\nu\}$ is of the form $\Psi_{\nu(\bullet)}$, where $\nu : [0,\beta] \mapsto \Omega_\Lambda$ is a piecewise constant classical spin configuration valued function. In fact relations (2.28) and (2.29) imply that, whatever are the realizations of Poisson processes $\xi = \{\xi_{ij}, \xi_i\}$ the only compatible $\nu \sim \xi$ are constant configurations $\nu(\bullet) \equiv const$. Furthermore, an arrival of K_{ij} at time t imposes and additional constraint $\nu_i(t) = \nu_j(t)$, whereas an arrival of K_i imposes an additional constraint $\nu_i(t) = 1$. It is convenient to explore (2.25) in terms of the following graphical representation (see Figure 1 below):

To each site $i \in \Lambda$ we attach a time interval $\mathbb{S}_\beta \stackrel{\Delta}{=} [0,\beta]$. In order to distinguish between intervals attached to different sites we use notation $\mathbb{S}^i_\beta$. Points on $\mathbb{S}^i_\beta$ labeled as (i,t). An arrival of ξ_{ij} at time t is visualized as a link between (i,t) and (j,t). An arrival of ξ_i at time t puts a $*$-mark at (i,t). It is also convenient to think about all $*$-marks being linked (wired) to some ghost site $\mathfrak{g}$. Two intervals $\mathbb{S}^i_\beta$ and $\mathbb{S}^j_\beta$ are said to be connected if $\xi_{ij} \neq \emptyset$. Thus, any realization of $\{\xi_{ij}\}$ splits

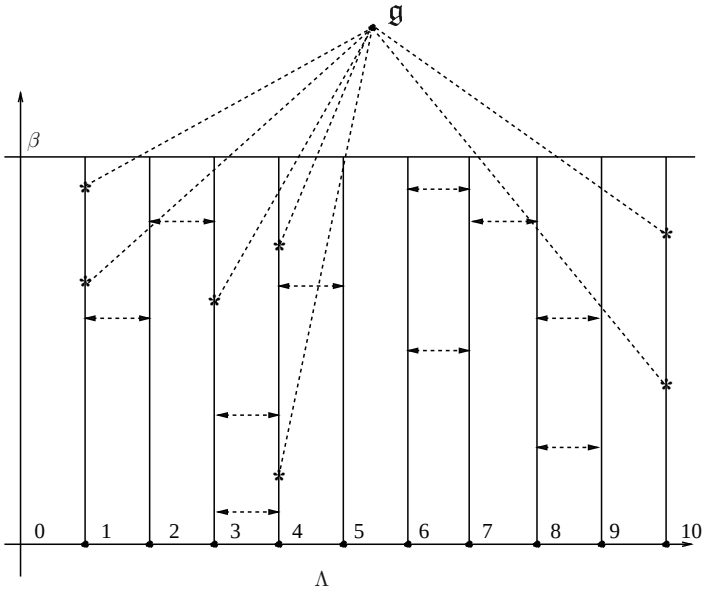


Fig. 1. The box Λ is split into three connected components, $\mathcal{C}_1 = \{1,2,3,4,5\} \times \mathbb{S}_\beta$, $\mathcal{C}_2 = \{6,7,8,9\} \times \mathbb{S}_\beta$ and $\mathcal{C}_3 = \{10\} \times \mathbb{S}_\beta$. Components $\mathcal{C}_1$ and $\mathcal{C}_3$ are wired, whereas $\mathcal{C}_2$ is free

$$\Lambda \times \mathbb{S}_\beta \,=\, \bigcup_{i\in\Lambda} \mathbb{S}^i_\beta \,=\, \cup \mathcal{C}_l$$

into the union of maximal connected components. Of course, each $\mathcal{C}_l$ above corresponds to a subset $\mathcal{A}_l$ of Λ,

$$\mathcal{C}_l \,=\, \bigcup_{i\in\mathcal{A}_l} \mathbb{S}^i_\beta.$$

A component $\mathcal{C}_l$ is said to be wired if $\xi_i \neq \emptyset$ for some $i \in \mathcal{A}_l$. It is convenient to link all wired components into one connected component. Given a realization $\xi = \{\xi_{ij}, \xi_i\}$ of all Poisson processes of arrivals of operators K_{ij} and K_i let $\#_{\mathsf{w}}(\xi)$ be the number of all maximal connected components $\mathcal{C}_l$ which are *not* wired to the ghost site $\mathfrak{g}$. Then the number of (constant) classical trajectories which satisfy (2.28) and (2.29) is precisely $2^{\#_{\mathsf{w}}(\xi)}$. For each such trajectory $\nu(\bullet) \equiv \nu$,

$$\prod_{t\in\xi} \langle \Psi_\nu | K_{\mathfrak{l}(t)} | \Psi_\nu \rangle \,=\, 1.$$

Consequently, let $\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}$ be the (Poisson) distribution of ξ. Then, (2.25) implies,

$$\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right) \,=\, \mathrm{e}^{\beta(\sum_{(i,j)} J_{ij} + \sum_i h)} \mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(2^{\#_{\mathsf{w}}(\xi)}\right). \tag{2.30}$$

Define a new measure $\widetilde{\mathbb{P}}^{\mathbf{J},h}_{\beta,\Lambda}$ on trajectories of point processes ξ,

$$\widetilde{\mathbb{P}}^{\mathbf{J},h}_{\beta,\Lambda}(\mathrm{d}\xi) \,=\, \frac{2^{\#_{\mathsf{w}}(\xi)} \mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}(\mathrm{d}\xi)}{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(2^{\#_{\mathsf{w}}(\xi)}\right)}. \tag{2.31}$$

$\widetilde{\mathbb{P}}^{\mathbf{J},h}_{\beta,\Lambda}$ is called FK or random cluster measure. Using (2.11) and (2.27) we arrive to the following stochastic geometric representation of classical expectations,

$$\mu^{\beta,h}_\Lambda(\nu_i) \,=\, \widetilde{\mathbb{P}}^{\mathbf{J},h}_{\beta,\Lambda}(i \longleftrightarrow \mathfrak{g}) \qquad \text{and} \qquad \mu^{\beta,h}_\Lambda(\nu_i\nu_j) \,=\, \widetilde{\mathbb{P}}^{\mathbf{J},h}_{\beta,\Lambda}(i \longleftrightarrow j)\,, \tag{2.32}$$

where the event $\{i \longleftrightarrow \mathfrak{g}\}$ means that the connected component of $\mathbb{S}^i_\beta$ is wired, whereas $\{i \longleftrightarrow j\}$ means that $\mathbb{S}^i_\beta$ and $\mathbb{S}^j_\beta$ belong to the same connected component (including the case when $\{i \longleftrightarrow \mathfrak{g}\} \cap \{j \longleftrightarrow \mathfrak{g}\}$).

2.4 Random Current Representation

In its turn classical RC representation corresponds to the decomposition of the Hamiltonian $\mathcal{H}_\Lambda$ in (2.9) as,

$$-\mathcal{H}_\Lambda \,=\, \sum_{(i,j)} J_{ij}\hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}} + \sum_i h\hat{\sigma}_i^{\mathsf{z}}.$$

The trick is to compute matrix elements of $\mathrm{e}^{-\beta\mathcal{H}_\Lambda}$ in the x-basis, which is defined as follows: With $\psi_{\pm 1}$ being defined as in (2.5), set

$$\phi_{\pm 1} \,=\, \frac{1}{\sqrt{2}}\,(\psi_1 \pm \psi_{-1})\,. \tag{2.33}$$

Clearly $\{\phi_{-1}, \phi_1\}$ is an orthonormal basis of $\mathbb{R}^2$. To a given classical x-configuration $\vartheta \in \Omega_\Lambda$ one corresponds the vector,

$$\Phi_\vartheta \,=\, \otimes_{i\in\Lambda}\phi_{\vartheta_i}. \tag{2.34}$$

The collection $\{\Phi_\vartheta\}$ is an orthonormal basis of $\mathbb{X}_\Lambda$. In the x-basis Pauli matrix $\hat{\sigma}^{\mathsf{z}}$ looks like

$$\hat{\sigma}^{\mathsf{z}} \,=\, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{or} \quad \hat{\sigma}^{\mathsf{z}}\phi_{\pm 1} = \phi_{\mp 1}. \tag{2.35}$$

Thus in the x-basis $\hat{\sigma}^{\mathsf{z}}$ is just a spin-flip operator. As in (2.7) the action of $\hat{\sigma}_i^{\mathsf{z}}$ on Φ_ϑ is given by

$$\hat{\sigma}_i^{\mathsf{z}}\Phi_\vartheta \,=\, \phi_{\vartheta_1} \otimes \cdots \otimes \hat{\sigma}^{\mathsf{z}}\phi_{\vartheta_i} \otimes \ldots.$$

In other words, $\hat{\sigma}_i^{\mathsf{z}}$ flips i-th component of Φ_ϑ.

In the language of Subsection 2.2, we are dealing with independent Poisson processes ξ_{ij} of arrivals of operators $K_{ij} \stackrel{\Delta}{=} \hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}}$ with intensities J_{ij} and with independent Poisson processes ξ_i of arrivals of operators $K_i \stackrel{\Delta}{=} \hat{\sigma}_i^{\mathsf{z}}$ with intensities h each. We continue to employ the running notation $\mathbb{P}_{\beta,\Lambda}^{\mathbf{J},h}$ for the distribution of ξ.

Let $\vartheta, \vartheta' \in \Omega_\Lambda$ be two classical x-configurations and let, as before, Φ_ϑ and $\Phi_{\vartheta'}$ be the corresponding elements of the x-basis of $\mathbb{X}_\Lambda$. Then,

$$\langle\Psi_\vartheta|K_{ij}|\Psi_{\nu'}\rangle \,=\, \delta_{\{\vartheta'=\hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}}\vartheta\}} \quad \text{and} \quad \langle\Psi_\vartheta|K_i|\Psi_{\nu'}\rangle \,=\, \delta_{\{\vartheta'=\hat{\sigma}_i^{\mathsf{z}}\vartheta\}}. \tag{2.36}$$

In other words, each arrival of operator K_{ij} enforces a simultaneous flip of i-th and j-th coordinate of Φ, and each arrival of operator K_i enforces a flip of i-th coordinate of Φ. Therefore, given a realization of

ξ, compatible space-time configurations $\Phi(\cdot) \sim \xi$ are deterministically recovered from the initial value $\Phi(0)$. Therefore, there are exactly $2^{|\Lambda|}$ compatible configurations for each realization of ξ.

Consider now the representation of the trace in (2.26). Clearly a space-time configuration $\Phi(\cdot)$ contributes only if $\Phi(0) = \Phi(\beta)$. In view of the above description of action of operators K_{ij} and K_i, this obviously imposes a restriction on admissible realization of ξ: Namely, there are trajectories $\Phi \sim \xi$ with $\Phi(0) = \Phi(\beta)$ if and only if ξ flips each coordinate $i \in \Lambda$ *even* number of times.

With a slight abuse of notation let ξ_{ij} and ξ_i also denote the number of arrivals of K_{ij}, respectively K_i on the interval $[0, \beta]$. In this way ξ will be called random currents. The total current through $i \in \Lambda$ is $\xi[i] = \sum_j \xi_{ij} + \xi_i$ and the total current through the ghost site $\mathfrak{g}$ is $\xi[\mathfrak{g}] = \sum_i \xi_i$. The boundary of a current is,

$$\partial\xi \stackrel{\Delta}{=} \{u \in \Lambda \cup \mathfrak{g} \,:\, \xi[u] \text{ is odd}\}\,. \tag{2.37}$$

If $\partial\xi = \emptyset$, then all of $2^{|\Lambda|}$ compatible configurations $\Psi(\bullet) \sim \xi$ satisfy $\Phi(0) = \Phi(\beta)$, otherwise (if $\partial\xi \neq \emptyset$) none of them is periodic. Consequently, (2.26) implies,

$$\frac{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)}{\mathrm{e}^{\beta(\sum_{(i,j)} J_{ij} + \sum_i h)}} \;=\; 2^{|\Lambda|}\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\xi = \emptyset\right). \tag{2.38}$$

The following representation of one and two point functions is now almost straightforward,

$$\mu^{\beta,h}_\Lambda(\nu_i) \;=\; \frac{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\xi = \{i,\mathfrak{g}\}\right)}{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\xi = \emptyset\right)} \text{ and } \mu^{\beta,h}_\Lambda(\nu_i\nu_j) \;=\; \frac{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\xi = \{i,j\}\right)}{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\xi = \emptyset\right)} \tag{2.39}$$

Switching Lemma

Let ξ and η be two independent random currents distributed according to the product Poisson measure $\mathbb{P}^{\mathbf{J},h}_\beta$ each. We continue to $\mathbb{P}^{\mathbf{J},h}_\beta$ to denote the product measure. Then, for every i, j and for every subset $A \subseteq \Lambda \cup \mathfrak{g}$,

$$\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\xi = \{i,j\}\right)\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\eta = A\right) = \mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\big(\partial\xi = \emptyset; \partial\eta = A\Delta\{i,j\}\,; i \stackrel{\xi+\eta}{\longleftrightarrow} j\big), \tag{2.40}$$

where the event $\left\{i \stackrel{\xi+\eta}{\longleftrightarrow} j\right\}$ means that there exists a path of bonds $b \in \mathcal{E}$ from i to j with $\xi(b) + \eta(b) > 0$.

We refer to [1] for a proof of (2.40). In view of (2.39) an immediate consequence is the following representation of the truncated two-point function:

$$\mu_{\Lambda}^{\beta,h}(\nu_i\nu_j) - \mu_{\Lambda}^{\beta,h}(\nu_i)\,\mu_{\Lambda}^{\beta,h}(\nu_j) = \frac{\mathbb{P}_{\beta,\Lambda}^{\mathbf{J},h}\Big(\partial\xi = \emptyset; \partial\nu = \{i,j\}\,; i \overset{\xi+\nu}{\longleftrightarrow} \mathfrak{g}\Big)}{\mathbb{P}_{\beta,\Lambda}^{\mathbf{J},h}\left(\partial\xi = \emptyset; \partial\eta = \emptyset\right)}. \tag{2.41}$$

Exponential Decay of Two-point Functions at Non-zero Magnetic Fields

Representation (2.41) and similar formulas pave the way for a stochastic geometric interpretation of semi-invariants and give rise to a useful intuition. As an example let us show how (2.41) implies that classical Ising truncated two-point functions always have non-zero exponential rate of decay once $h \neq 0$. The argument below was developed together with Roberto Fernandez and Yvan Velenik some time ago. As it was pointed out by Yvan, a conventional proof could be found in [12].

Let $\kappa = \xi + \eta$ be the combined current. Any realization of κ splits Λ into a disjoint union of maximal connected components: as before we say that i and j are connected if there exists a chain of bonds b leading from i to j with $\kappa(b) > 0$ on each bond. Clearly $\{\partial\xi = \emptyset; \partial\eta = \{i,j\}\}$ implies that $\partial\kappa = \{i,j\}$ and, in particular that i and j are connected in κ or, in other words, that i and j belong to the same connected component C of κ. If R is the range of interaction, then $|C| \geq |i-j|/R$, as soon as we impose an additional constraint $\left\{C \overset{\kappa}{\longleftrightarrow} \mathfrak{g}\right\}$. It is almost obvious now why (2.41) implies exponential decay: one should pay a fixed price to disconnect each site $l \in C$ from the ghost site $\mathfrak{g}$, see Figure 2.

It remains to make the last remark precise. For any connected set $C \subset \Lambda$ define $\#\left(\mathcal{E}(C, \Lambda \setminus C)\right)$ as the number of edges in $\mathcal{E}(C, \Lambda \setminus C)$, where the latter is the set of edges b with $J_b > 0$, which have one endpoint in C and another in $\Lambda \setminus C$. The probability $p(C, \Lambda \setminus C)$ that none of the processes $\kappa_b;\ b \in \mathcal{E}(C, \Lambda \setminus C)$ arrives on the interval $[0, \beta]$ is

$$p(C, \Lambda \setminus C) \;=\; \exp\Big\{-2\beta \sum_{b \in \mathcal{E}(C, \Lambda\setminus C)} J_b\Big\}.$$

Given a connected set C and $i, j \in C$, define the following event

$$\mathcal{A}_{ij}(C) \;=\; \Big\{\partial\xi_C = \emptyset; \partial\eta_C = \{i,j\}\,;\ C \text{ is connected in } \kappa_C;\ C \overset{\kappa_C}{\longleftrightarrow} \mathfrak{g}\Big\},$$

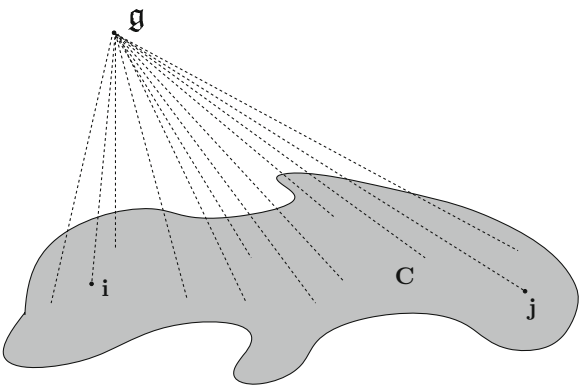


Fig. 2. Each site l inside the connected component C has a chance to be connected by a direct non-zero current to the ghost site $\mathfrak{g}$

where ξ_C, η_C (respectively $\xi_{\Lambda\setminus C}, \eta_{\Lambda\setminus C}$) and κ_C (respectively $\kappa_{\Lambda\setminus C}$) are the restrictions of the corresponding processes to the bonds with either both end-points at C ($\Lambda \setminus C$) or with one end-point at C ($\Lambda \setminus C$) and another end-point being $\mathfrak{g}$. In this notation the expression in the numerator in (2.41)

$$\sum_{C\,\text{connected}} \mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\xi_{\Lambda\setminus C}=\emptyset;\partial\eta_{\Lambda\setminus C}=\emptyset\right)\cdot p(C,\Lambda\setminus C)\cdot \mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\mathcal{A}_{ij}(C)\right). \tag{2.42}$$

On the other hand, the denominator in (2.41) is certainly bounded below by

$$\sum_{C\,\text{connected}} \mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\partial\xi_{\Lambda\setminus C}=\emptyset;\partial\eta_{\Lambda\setminus C}=\emptyset\right)\cdot p(C,\Lambda\setminus C)\cdot \mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\mathcal{A}^{\mathrm{e}}_{ij}(C)\right), \tag{2.43}$$

where the event

$$\mathcal{A}^{\mathrm{e}}_{ij}(C) \;=\; \{\partial\xi_C=\emptyset;\partial\eta_C=\emptyset;\; C \text{ is connected in } \kappa_C\}.$$

We claim that there exist two positive constants c_1 and c_2 which depend on β, h (but not on the range R of the interaction, the dimension of the lattice, connected C and $\{i,j\}\subseteq C$), such that,

$$\frac{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\mathcal{A}_{ij}(C)\right)}{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\left(\mathcal{A}^{\mathrm{e}}_{ij}(C)\right)} \;\le\; c_1\mathrm{e}^{-c_2|i-j|/R}. \tag{2.44}$$

Indeed, for each current η_C with $\partial\eta = \{i,j\}$ and $\mathfrak{g} \overset{\eta_C}{\longleftrightarrow\!\!\!\!\!/\;\;} C$, we may construct a family of currents

$$\{\eta_C\}^{\mathrm{e}} = \Big\{\eta + (2r_i+1)\delta_{(i,\mathfrak{g})} + (2r_j+1)\delta_{(j,\mathfrak{g})} + \sum_{k\in C\setminus\{i,j\}} 2r_k\delta_{(k,\mathfrak{g})}\Big\}_{r_l=0,1,\ldots \text{for } l\in C}.$$

Thus, the family $\{\eta_C\}^{\mathrm{e}}$ is generated by tuples $\underline{r} = \{r_l\}_{l\in\Lambda}$ of non-negative integers. Evidently,

$$\{\eta_C\}^{\mathrm{e}} \cap \{\eta'_C\}^{\mathrm{e}} = \emptyset,$$

whenever $\eta_C \neq \eta'_C$. Furthermore, $(\xi_C, \eta_C) \in \mathcal{A}_{ij}(C) \Rightarrow \xi_C \times \{\eta_C\}^{\mathrm{e}} \subseteq \mathcal{A}^{\mathrm{e}}_{ij}(C)$. However, for such η_C,

$$\frac{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}(\{\eta_C\}^{\mathrm{e}})}{\mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}(\eta_C)} = (\sinh(\beta h))^2 \cdot (\cosh(\beta h))^{|C|-2},$$

and (2.44) follows.

3 Quantum Ising Models in Transverse Field

Quantum Ising Hamiltonian in transverse field λ is given by

$$-\mathcal{H}_\Lambda = \sum_{(i,j)} J_{ij}\hat{\sigma}^{\mathsf{z}}_i\hat{\sigma}^{\mathsf{x}}_j + h\sum_i \hat{\sigma}^{\mathsf{z}}_i + \lambda\sum_i \hat{\sigma}^{\mathsf{x}}_i, \tag{3.1}$$

where $\lambda \geq 0$, and (in the z-basis),

$$\hat{\sigma}^{\mathsf{z}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad \text{and} \quad \hat{\sigma}^{\mathsf{x}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.2}$$

Since matrices $\hat{\sigma}^{\mathsf{x}}$ and $\hat{\sigma}^{\mathsf{z}}$ do not commute, as soon as the strength of the transverse field $\lambda > 0$, the operator $\mathcal{H}_\Lambda$ does not have diagonal form neither in z-basis (2.6), nor in the x-basis (2.34). Nevertheless, the analog of Lie-Trotter product formula still holds,

$$\mathrm{e}^{-\beta\mathcal{H}_\Lambda} = \lim_{\Delta\to 0}\Big(\prod_{(i,j)} \mathrm{e}^{\Delta J_{ij}\hat{\sigma}^{\mathsf{z}}_i\hat{\sigma}^{\mathsf{z}}_j} \prod_i \mathrm{e}^{\Delta h\hat{\sigma}^{\mathsf{z}}_i} \prod_i \mathrm{e}^{\Delta\lambda\hat{\sigma}^{\mathsf{x}}_i}\Big)^{\beta/\Delta}. \tag{3.3}$$

As in the classical case various choices of bases and of decomposition of $\mathcal{H}_\Lambda$ lead to different stochastic geometric representations of the model.

3.1 FK Representation

As in the classical case the traces are computed in the z-basis. As for the decomposition represent $-\mathcal{H}_\Lambda$ as

$$-\Big(\sum_{(i,j)} J_{ij} + \sum_i h + \sum_i \lambda\Big)\mathrm{I} + \sum_{(i,j)} 2J_{ij}\frac{\mathrm{I}+\hat\sigma_i^{\mathsf z}\hat\sigma_j^{\mathsf z}}{2} + \sum_i 2h\frac{\mathrm{I}+\hat\sigma_i^{\mathsf z}}{2} + \sum_i \lambda(\hat\sigma_i^{\mathsf x}+\mathrm{I}).$$

In the language of Subsection 2.2 we are dealing with Poisson process ξ of arrivals on the interval $[0,\beta]$ of the following type of operators:

- Operators $K_{ij} = \left(\hat\sigma_i^{\mathsf z}\hat\sigma_j^{\mathsf z} + \mathrm{I}\right)/2$ which arrive with intensities $2J_{ij}$. We shall call these processes *links* and denote them as ξ_{ij}.
- Operators $K_i^h = (\hat\sigma_i^{\mathsf z} + \mathrm{I})/2$ which arrive with intensities $2h$. We shall call these processes *links to* $\mathfrak{g}$ and denote them as ξ_i^h.
- Operators $K_i^\lambda = \hat\sigma_i^{\mathsf x} + \mathrm{I}$ which arrive with intensities λ. We shall call these processes *holes* and denote them as ξ_i^λ.

As in Subsection 2.3 piece-wise constant functions $\Psi : [0,\beta] \mapsto \{\Psi_\nu\}$ are labeled by piece-wise constant classical trajectories $\nu : [0,\beta] \mapsto \Omega_\Lambda$. Given a realization $\xi = \left\{\xi_{ij}, \xi_i^h, \xi_i^\lambda\right\}$ let us try to describe the family of compatible trajectories $\nu \sim \xi$.

1. Since $\langle\Psi_\nu|K_{ij}|\Psi_{\nu'}\rangle = \delta_{\{\nu=\nu'\}}\delta_{\{\nu_i=\nu_j\}}$, an arrival of an (i,j)-link at time t imposes the constraint $\nu(t,i) = \nu(t,j)$.
2. Since $\langle\Psi_\nu|K_i^h|\Psi_{\nu'}\rangle = \delta_{\{\nu=\nu'\}}\delta_{\{\nu_i=1\}}$, an arrival of an $(i,\mathfrak{g})$-link at time t imposes the constraint $\nu(t,i) = 1$.
3. Since
$$\langle\Psi_\nu|K_i^\lambda|\Psi_{\nu'}\rangle = \delta_{\left\{\nu_j=\nu'_j \quad \text{for all } j\neq i\right\}},$$
an arrival of an i-hole at time t *enables* a flip of i-th coordinate of $\nu(t,\cdot)$.

Thus, contrary to the classical situation considered in Subsection 2.3, compatible configurations $\nu \sim \xi$ are permitted to have jumps at arrival times of ξ^λ. It is convenient to visualize compatible *periodic* $\nu(\cdot)$ as follows (see Figure 3): For each $i \in \Lambda$ the process of holes ξ_i^λ splits the circle $\mathbb{S}^i_\beta$ (which is the interval $i \times [0,\beta]$ with the end-points $(i,0)$ and (i,β) identified) into a disjoint union of connected intervals. Two such intervals $i \times I \subseteq \mathbb{S}^i_\beta$ and $j \times J \subseteq \mathbb{S}^j_\beta$ are said to be connected in ξ if there is an arrival of ξ_{ij} at a time $t \in I \cap J$. A maximal connected cluster $\cup_l \{i_l \times I_l\}$ (with i_l-s being not necessarily different, but with $\{i_l \times I_l\} \cap \{i_m \times I_m\} = \emptyset$ whenever $l \neq m$) is said to be connected to the ghost site $\mathfrak{g}$ if for some for some i_l a process $\xi^h_{i_l}$ arrives at $t \in I_l$. Otherwise

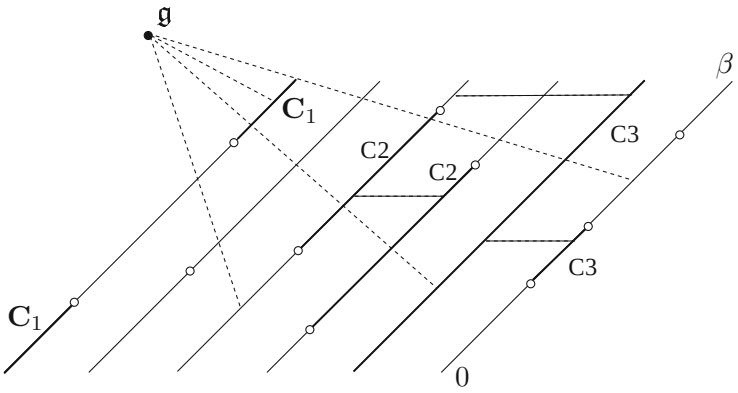


Fig. 3. Configurations with periodic boundary conditions $\nu(i,0) = \nu(i,\beta)$. Connected components $\mathbf{C}_1$ and $\mathbf{C}_3$ are linked to $\mathfrak{g}$ and hence $\nu \equiv 1$ on them. Connected component $\mathbf{C}_2$ is "free" and hence one can colour it in either of ± 1 colours.

such maximal connected cluster is called free. Define $\#_{\mathsf{w}}(\xi)$ to be the number of maximal free connected clusters of ξ. Then, using $\mathbb{P}^{\mathbf{J},h,\lambda}_{\beta,\Lambda}$ for the reference product distribution of independent Poisson processes ξ, we arrive to the following quantum version of the FK representation (2.30) of the trace,

$$\mathsf{Tr}\big(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\big) \;=\; \mathrm{e}^{\beta(\sum_{(i,j)} J_{ij} + \sum_i h + \sum_i \lambda)} \mathbb{P}^{\mathbf{J},h}_{\beta,\Lambda}\big(2^{\#_{\mathsf{w}}(\xi)}\big). \tag{3.4}$$

As in the classical case define a new measure $\widetilde{\mathbb{P}}^{\mathbf{J},h,\lambda}_{\beta,\Lambda}$ on trajectories of point processes ξ,

$$\widetilde{\mathbb{P}}^{\mathbf{J},h,\lambda}_{\beta,\Lambda}(\mathrm{d}\xi) \;=\; \frac{2^{\#_{\mathsf{w}}(\xi)} \mathbb{P}^{\mathbf{J},h,\lambda}_{\beta,\Lambda}(\mathrm{d}\xi)}{\mathbb{P}^{\mathbf{J},h,\lambda}_{\beta,\Lambda}\big(2^{\#_{\mathsf{w}}(\xi)}\big)}. \tag{3.5}$$

Once again, using (2.11) and (2.27) we arrive to the following stochastic geometric representation of expectations,

$$\frac{\mathsf{Tr}\left(\hat{\sigma}^{\mathsf{z}}_i \mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)}{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)} \;=\; \widetilde{\mathbb{P}}^{\mathbf{J},h,\lambda}_{\beta,\Lambda}\left((i,0) \longleftrightarrow \mathfrak{g}\right) \tag{3.6}$$

where the event $\{(i,0) \longleftrightarrow \mathfrak{g}\}$ means that the $\mathbb{S}^i_\beta$ interval containing $(i,0)$ belongs to a cluster which is connected to $\mathfrak{g}$. Similarly,

$$\frac{\mathsf{Tr}\big(\hat{\sigma}^{\mathsf{z}}_i \hat{\sigma}^{\mathsf{z}}_j \mathrm{e}^{-\beta\mathcal{H}_\Lambda}\big)}{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)} \;=\; \widetilde{\mathbb{P}}^{\mathbf{J},h,\lambda}_{\beta,\Lambda}\left((i,0) \longleftrightarrow (j,0)\right), \tag{3.7}$$

where the event $\{(i,0) \longleftrightarrow (j,0)\}$ means that the corresponding $\mathbb{S}^i_\beta$ and $\mathbb{S}^j_\beta$ intervals belong to the same connected cluster.

Ground States, Matrix and Reduced Density Matrix Elements Let us fix a finite graph $(\Lambda, \mathcal{E})$, coupling constants $\mathbf{J}$ and $\lambda \geq 0$. In order to facilitate the notation we shall set magnetic filed in z-direction to zero, $h = 0$. For each $\beta \in \mathbb{R}$, z-matrix elements $\rho^{\mathsf{z}}_\beta(\nu, \nu')$ are defined via,

$$\rho^{\mathsf{z}}_\beta(\nu,\nu') \;=\; \frac{\langle \Psi_\nu | \mathrm{e}^{-\beta \mathcal{H}_\Lambda} | \Psi_{\nu'} \rangle}{\mathsf{Tr}\left(\mathrm{e}^{-\beta \mathcal{H}_\Lambda}\right)}. \tag{3.8}$$

In order to derive an appropriate expression in terms of Poisson arrival measures $\mathbb{P}^{\mathbf{J},\lambda}_{\beta,\Lambda}$ or in terms of the FK measures $\mathbb{P}^{\mathbf{J},\lambda}_{\beta,\Lambda}$ we should introduce a modification of the notion of connected components of ξ. Originally, those were defined as unions of sub-intervals of $\mathbb{S}_\beta$. However, in the computation of matrix elements we, obviously, do not impose periodicity conditions. In the sequel, given a subset $A \subset \Lambda$ and a configuration ξ let ξ_A be obtained from ξ via adding holes at all the points $(i,0) = (i,\beta)$ with $i \in A$. One can think about ξ_A in terms of slitting the A-part of ξ along $t = 0$.

Any piece-wise trajectory $\nu : [0,\beta] \to \Omega_\Lambda$ which contributes to the numerator in (3.8) satisfies boundary conditions,

$$\nu(i,0) \;=\; \nu_i \quad \text{and} \quad \nu(i,\beta) \;=\; \nu'_i \quad \forall\, i \in \Lambda.$$

As a result, realizations of ξ which place points (i,T) and (j,S) (with $i,j \in \Lambda$ and $T,S = 0\,\text{or}\,\beta$) with $\nu(i,T) \neq \nu(j,S)$ into same connected components of the slit configuration ξ_Λ do not have compatible trajectories at all. Let us say that $\xi_\Lambda \sim \{\nu,\nu'\}$, if the latter does not happen. If $\xi_\Lambda \sim \{\nu,\nu'\}$, then the set of all ξ-compatible trajectories, which contribute to the denominator in (3.8) is constructed in the following way: Each connected cluster of ξ_Λ whose closure hits either $t = 0$ or $t = \beta$ layers inherits the z-spin value from ν or ν'. On the other hand, each interior cluster of ξ_Λ or, alternatively each cluster of ξ which does not contain points with $0 = \beta$ time coordinates, could be still coloured into ± 1. Clusters of ξ which are not interior are called boundary. Thus, if we use $\#_0(\xi)$ and $\#_\partial(\xi) = \#(\xi) - \#_0(\xi)$ for the number of interior (respectively boundary) clusters of ξ,

$$\rho^{\mathsf{z}}_\beta(\nu,\nu') = \frac{\mathbb{P}^{\mathbf{J},\lambda}_{\beta,\Lambda}\left(\xi_\Lambda \sim \{\nu,\nu'\}\;;\, 2^{\#_0(\xi)}\right)}{\mathbb{P}^{\mathbf{J},\lambda}_{\beta,\Lambda}\left(2^{\#(\xi)}\right)} = \widetilde{\mathbb{P}}^{\mathbf{J},\lambda}_{\beta,\Lambda}\big(\xi_\Lambda \sim \big\{\nu,\nu'\big\}\;;\, 2^{-\#_\partial(\xi)}\big). \tag{3.9}$$

For each $\lambda > 0$ there exist non-trivial limits $\mathbb{P}^{\mathbf{J},\lambda}_{\infty,\Lambda}$ and $\widetilde{\mathbb{P}}^{\mathbf{J},\lambda}_{\infty,\Lambda}$ as $\beta \to \infty$. These measures could be constructed directly: $\mathbb{P}^{\mathbf{J},\lambda}_{\infty,\Lambda}$ is just the distribution of Poisson processes of arrival ξ on $\mathbb{R}$. Connected components of ξ are understood now as linked sub-intervals of $\mathbb{R}$ over various spatial coordinates $i \in \Lambda$. The FK measure $\widetilde{\mathbb{P}}^{\mathbf{J},\lambda}_{\infty,\Lambda}$ is then constructed via modification of $\mathbb{P}^{\mathbf{J},\lambda}_{\infty,\Lambda}$ by the $2^{\#(\xi)}$ factor (as a limiting procedure, of course). Boundary clusters of the slit configuration ξ_Λ are coloured in this way according to ν just above the $t = 0$ layer and according to ν' just below it. If we slit along all of Λ, then the compatibility condition $\xi_\Lambda \sim \{\nu, \nu'\}$ decouples into $\{\xi_+ \sim \nu\} \cap \{\xi_- \sim \nu'\}$ for the upper and lower halves ξ_+ and ξ_- of configuration ξ. At this point it makes sense to introduce Poisson $\mathbb{P}^{\mathbf{J},\lambda,+}_{\infty,\Lambda}$ and, accordingly, FK $\widetilde{\mathbb{P}}^{\mathbf{J},\lambda,+}_{\infty,\Lambda}$ measures for arrival processes on $\mathbb{R}_+$. It is straightforward now to check that matrix elements $\rho^{\mathsf{z}}_{\infty}(\nu,\nu') = \langle \Psi_\nu | \Psi \rangle \langle \Psi | \Psi_{\nu'} \rangle$, which are generated by projections of the ground state Ψ of $\mathcal{H}_\Lambda$ are given by,

$$\rho^{\mathsf{z}}_{\infty}(\nu,\nu') \;=\; \widetilde{\mathbb{P}}^{\mathbf{J},\lambda}_{\infty,\Lambda}\big(\xi_\Lambda \sim \big\{\nu,\nu'\big\}\;;\, 2^{-\#_\partial(\xi)}\big). \tag{3.10}$$

In the notation just introduced above the latter expression equals to

$$\langle \Psi_\nu | \Psi \rangle \langle \Psi | \Psi_{\nu'} \rangle \;=\; \widetilde{\mathbb{P}}^{\mathbf{J},\lambda,+}_{\infty,\Lambda}\big(\xi \sim \nu;\, 2^{-\#_\partial(\xi)}\big)\widetilde{\mathbb{P}}^{\mathbf{J},\lambda,+}_{\infty,\Lambda}\big(\xi \sim \nu';\, 2^{-\#_\partial(\xi)}\big).$$

Similarly, for $A \subseteq \Lambda$ and $\theta, \theta' \in \{\pm 1\}^A$, the reduced density matrix entry $\rho^{\mathsf{z}}_{\infty,A}(\theta,\theta')$ is given by

$$\rho^{\mathsf{z}}_{\infty,A}(\theta,\theta') \;=\; \widetilde{\mathbb{P}}^{\mathbf{J},\lambda}_{\infty,\Lambda}\big(\xi_A \sim \big\{\theta,\theta'\big\}\;;\, 2^{-\#_{\partial,A}(\xi)}\big), \tag{3.11}$$

where the compatibility condition $\xi_A \sim \{\nu,\nu'\}$ for the slit configuration ξ_A is defined in the obvious way, and $\#_{\partial,A}(\xi)$ stands for the number of connected clusters of ξ which contain points $(0,i)$ with $i \in A$.

3.2 Random Current Representation

In order to derive an appropriate version of random current representation let us rewrite the Hamiltonian (3.1) as

$$-\Big(\sum_i \lambda\Big)\mathrm{I} + \sum_{(i,j)} J_{ij}\hat{\sigma}^{\mathsf{z}}_i\hat{\sigma}^{\mathsf{z}}_j + \sum_i h\hat{\sigma}^{\mathsf{z}}_i + \sum_i 2\lambda\frac{\hat{\sigma}^{\mathsf{x}}_i + \mathrm{I}}{2}.$$

As in the classical case the traces are going to be computed in the x-basis (2.34). Thus, in the language of Subsection 2.2 we are dealing

with Poisson process ξ of independent arrivals on $[0,\beta]$ of the following type of operators:

- Operators of simultaneous (ij)-flips $K_{ij}=\hat\sigma_i^{\mathsf{z}}\hat\sigma_j^{\mathsf{z}}$ which arrive with intensities J_{ij}. We shall denote the corresponding Poisson process ξ_{ij}.
- Operators of i-flips $K_i^h=\hat\sigma_i^{\mathsf{z}}$ which arrive with intensity h each. The corresponding Poisson processes are denoted as ξ_i^h.
- Operators $K_i^\lambda=(\hat\sigma^{\mathsf{x}}+\mathrm{I})/2$ which arrive with intensity 2λ each. The corresponding Poisson process is denoted ξ_i^λ. Since,

$$\langle\Phi_\vartheta|K_i^\lambda|\Phi_{\vartheta'}\rangle\;=\;\delta_{\{\vartheta=\vartheta'\}}\delta_{\{\vartheta_i=1\}},$$

  an arrival of ξ_i^λ at time t imposes the constraint $\vartheta(i,t)=1$ for every ξ-compatible classical piece-wise constant x-trajectory $\vartheta:[0,\beta]\mapsto\Omega_\Lambda$. We shall refer to ξ^λ as to processes of *marks*.

Accordingly, for a given realization of ξ compatible *periodic* piece-wise constant trajectories $\vartheta(\cdot)$ are characterized as follows:

1. Arrivals of ξ_{ij} and of ξ_i^h enforce simultaneous flips of i-th and j-th coordinates of ϑ, respectively of i-th coordinate of ϑ. These are the only jumps of $\vartheta(\cdot)$.
2. For each $i\in\Lambda$, $\vartheta(i,t)=1$ at all arrival times of ξ_i^λ.

Let us try to compute the number of ξ-compatible trajectories ϑ for a given realization ξ. It is natural to modify the notion of the boundary $\partial\xi$ as follows: For every $i\in\Lambda$ the process of marks ξ_i^λ splits the circle $\mathbb{S}_\beta^i$ into the disjoint union of intervals,

$$\mathbb{S}_\beta^i\setminus\xi_i^\lambda\;=\;\cup_{l=1}^{m(i)}J_l^{(i)}\;\stackrel{\Delta}{=}\;\cup_{l=1}^{m(i)}i_l\times I_l^{(i)}. \tag{3.12}$$

The number $m(i)$ of such disjoint intervals equals to 1 if $\xi_i^\lambda=0$ and to ξ_i^λ otherwise. Let us say that an interval $J_l^{(i)}$ in the decomposition (3.12) belongs to the boundary $\partial\xi$ if (see Figure 4) the total current through $J_l^{(i)}$

$$\xi[J_l^{(i)}]\;\stackrel{\Delta}{=}\;\sum_{j\in\Lambda\setminus i}\xi_{ij}(J_l^{(i)})\;+\;\xi_i^h(J_l^{(i)}),$$

is odd. Evidently, there are periodic compatible $\vartheta\sim\xi$ iff $\partial\xi=\emptyset$. In the later case, there is a unique compatible trajectory $\nu(i,\cdot)$ for every *marked* $i\in\Lambda$ such that $\xi_i^\lambda>0$ and, accordingly, there are precisely two compatible trajectories for every *unmarked* i with $\xi_i^\lambda=0$. Let $\#_{\mathsf{m}}(\xi)\;=\;\#\left\{i:\xi_i^\lambda=0\right\}$ be the total number of unmarked intervals $[0,\beta]$. By the general trace formula (2.26),

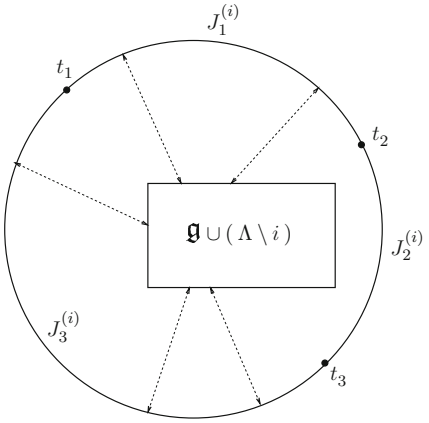


Fig. 4. The arrival times of the process of marks ξ_i^λ are t_1, t_2 and t_3. Accordingly, $\mathbb{S}_\beta^i$ is split into three marked intervals $J_1^{(i)}, J_2^{(i)}$ and $J_3^{(i)}$. The total number of arrivals of flips on $J_3^{(i)}$ equals to three, hence $J_3^{(i)} \subseteq \partial\xi$.

$$\frac{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)}{\mathrm{e}^{\beta(\sum_{(i,j)} J_{ij} + \sum_i h + \sum_i \lambda)}} \;=\; \mathbb{P}_\beta^{\mathbf{J},h,\lambda}\big(2^{\#_{\mathsf{m}}(\xi)}\,;\,\partial\xi = \emptyset\big). \tag{3.13}$$

Thus, contrary to what happened in the the classical case, one should modify the reference (Poisson) measure. Define,

$$\widetilde{\mathbb{P}}_{\beta,\Lambda}^{\mathbf{J},\lambda,h}(\mathrm{d}\xi) \;=\; \frac{2^{\#_{\mathsf{m}}(\xi)}\mathbb{P}_{\beta,\Lambda}^{\mathbf{J},\lambda,h}(\mathrm{d}\xi)}{\mathbb{P}_{\beta,\Lambda}^{\mathbf{J},\lambda,h}\left(2^{\#_{\mathsf{m}}(\xi)}\right)}.$$

Then, as in the classical case, the following random current representation of one and two point functions hold: Let $J(i,t)$ be the marked interval containing (i,t). Then,

$$\frac{\mathsf{Tr}\left(\hat{\sigma}_i^{\mathsf{z}}\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)}{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)} \;=\; \frac{\widetilde{\mathbb{P}}_{\beta,\Lambda}^{\mathbf{J},\lambda,h}\left(\partial\xi = J(i,0) \cup \mathfrak{g}\right)}{\widetilde{\mathbb{P}}_{\beta,\Lambda}^{\mathbf{J},\lambda,h}\left(\partial\xi = \emptyset\right)} \tag{3.14}$$

and, similarly,

$$\frac{\mathsf{Tr}\left(\hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}}\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)}{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}_\Lambda}\right)} \;=\; \frac{\widetilde{\mathbb{P}}_{\beta,\Lambda}^{\mathbf{J},h,\lambda}\left(\partial\xi = J(i,0) \cup J(j,0)\right)}{\widetilde{\mathbb{P}}_{\beta,\Lambda}^{\mathbf{J},h,\lambda}\left(\partial\xi = \emptyset\right)} \tag{3.15}$$

It is, of course, a very natural question what should be a correct analog of the switching lemma in the quantum case. A closed form answer is still missing, but some aspects of this issue are discussed in [9].[1]

[1] Appropriate versions of switching lemma were recently derived by Crawford and Ioffe [10] and by Björnberg and Grimmett [5].

Ground States, Matrix and Reduced Density Matrix Elements Let us briefly sketch how matrix and reduced density matrix elements in the x-basis could be written using the RC representation. Again, in order to simplify the notation we shall consider only the case of $h = 0$, and, exactly as in the end of Subsection 3.1, we shall directly pass to the ground state limit $\beta \to \infty$. In the ground state we are dealing with processes of arrivals ξ on the whole real line $\mathbb{R}$. We use $\mathbb{P}^{\mathbf{J},\lambda}_{\infty,\Lambda}$ to denote the corresponding product measure. Evidently, $\xi^\lambda_i \neq \emptyset \, \forall i$ $\mathbb{P}^{\mathbf{J},\lambda}_{\infty,\Lambda}$-a.s. In other words, for each $i \in \Lambda$ the copy of the real line associated with i contains marks. As in the FK case, given ξ and a subset $A \subseteq \Lambda$, we use ξ_A to denote the slit configuration: except that now we view ξ_A as ξ with additional marks placed at time zero for each $i \in A$.

With such notation in mind we classify all marked intervals of ξ^λ and ξ^λ_A as follows:

1. Marked interval $i \times I$ of ξ^λ belong to $\mathcal{M}_0(\xi^\lambda)$ if $0 \in I$. Otherwise it belongs to $\mathcal{M}_{\text{ext}}(\xi^\lambda)$.
2. Marked intervals of the type $i \times (0,t)$ of ξ_A belong to $\mathcal{M}^+_0(\xi_A)$. Similarly, marked intervals of the type $i \times (-t,0)$ of ξ_A belong to $\mathcal{M}^-_0(\xi^\lambda_A)$.
3. All other marked intervals are ξ_A are also marked intervals of ξ and we classify them as $\mathcal{M}_0(\xi^\lambda_A)$ and $\mathcal{M}_{\text{ext}}(\xi^\lambda_A)$.

Accordingly, we define the boundaries $\partial_0\xi$, $\partial_{\text{ext}}\xi$, $\partial^+_0\xi_A$, $\partial^-_0\xi_A$, $\partial_0\xi_A$ and $\partial_{\text{ext}}\xi_A$ as e.g.,

$$\partial_0\xi \, = \, \big\{ i \times I \in \mathcal{M}_0(\xi^\lambda) \, : \, \xi[i \times I] \text{ is odd} \big\}.$$

Let us introduce the following conditional measure

$$\mathbb{M}^{\mathbf{J},\lambda}_{\infty,\Lambda} \, = \, \mathbb{P}^{\mathbf{J},\lambda}_{\infty,\Lambda} \left(\cdot | \partial_{\text{ext}}\xi = \emptyset \right).$$

Since Λ is finite the above definition can be easily turned into a meaningful one via an appropriate limiting procedure.

Let $\vartheta, \vartheta' \in \{\pm 1\}^\Lambda$ be two classical x-configurations, and let $\rho^{\mathsf{x}}_\infty(\vartheta, \vartheta')$ be the corresponding matrix element. From our interpretation of a mark in terms of a $+1$-spin enforcement at the corresponding space-time arrival point, it is apparent that that ξ contributes to $\rho^{\mathsf{x}}_\infty(\vartheta, \vartheta')$ iff the following event $\mathcal{E}_\pm(\vartheta, \vartheta') = \mathcal{E}_+(\vartheta) \cap \mathcal{E}_-(\vartheta')$ occurs:

1. Event $\mathcal{E}_+(\vartheta)$: For every $i \times I \in \mathcal{M}^+_0(\xi^\lambda_\Lambda)$, $i \times I \in \partial^+_0\xi_A$ iff ϑ_i is -1.
2. Event $\mathcal{E}_-(\vartheta')$: For every $i \times I \in \mathcal{M}^-_0(\xi^\lambda_\Lambda)$, $i \times I \in \partial^-_0\xi_A$ iff ϑ'_i is -1.

Then,

$$\rho^{\mathsf{x}}_{\infty}\left(\vartheta,\vartheta'\right) \;=\; \frac{\mathbb{M}^{\mathbf{J},\lambda}_{\infty,\Lambda}\left(\mathcal{E}_{\pm}(\vartheta,\vartheta')\right)}{\mathbb{M}^{\mathbf{J}}_{\infty,\Lambda},\lambda\left(\partial_0\xi=\emptyset\right)}.$$

In a similar fashion for $A\subseteq\Lambda$ and two classical x-configurations $\theta,\theta'\in\{\pm1\}^A$ define the event $\mathcal{E}^A_{\pm}(\theta,\theta')$ exactly as above, except that even/odd conditions on currents are restricted to intervals $i\times I$ from $\mathcal{M}^{\pm}_0(\xi^{\lambda}_A)$. Then, the (θ,θ') entry of the reduced density matrix is given by,

$$\rho^{\mathsf{x}}_{\infty,A}\left(\theta,\theta'\right) \;=\; \frac{\mathbb{M}^{\mathbf{J},\lambda}_{\infty,\Lambda}\left(\mathcal{E}^A_{\pm}(\theta,\theta');\partial_0\xi_A=\emptyset\right)}{\mathbb{M}^{\mathbf{J},\lambda}_{\infty,\Lambda}\left(\partial_0\xi=\emptyset\right)}.$$

4 Curie-Weiss Model and Erdős-Rényi Random Graphs

Classical Curie-Weiss mean-field Hamiltonian $\mathbf{H}^{\mathsf{CW}}_N$ is a function on $\Omega_N=\{\pm1\}^N$,

$$-\mathbf{H}^{\mathsf{CW}}_N(\nu) \;=\; \frac{1}{N}\sum_{(i,j)}\nu_i\nu_j, \tag{4.1}$$

where, as before, the summation is over all unordered pairs of $i\neq j$. In the language of Subsection 2.1, $\left\{\mathbf{H}^{\mathsf{CW}}_N(\nu)\right\}$ are eigenvalues of the quantum Hamiltonian $\mathcal{H}^{\mathsf{CW}}_N$,

$$\mathcal{H}^{\mathsf{CW}}_N\Psi_\nu \;=\; \mathbf{H}^{\mathsf{CW}}_N(\nu)\Psi_\nu, \quad \text{where} \quad -\mathcal{H}^{\mathsf{CW}}_N=\frac{1}{N}\sum_{(i,j)}\hat\sigma^{\mathsf{z}}_i\hat\sigma^{\mathsf{z}}_j.$$

Accordingly, for a given value of the inverse temperature β, the distribution of ν is,

$$\mu^{\beta}_N(\nu) \;=\; \frac{1}{\mathcal{Z}_N}\mathrm{e}^{-\beta\mathbf{H}^{\mathsf{CW}}_N(\nu)} \;=\; \frac{\langle\Psi_\nu|\mathrm{e}^{-\beta\mathcal{H}^{\mathsf{CW}}_N(\nu)}|\Psi_\nu\rangle}{\mathsf{Tr}\left(\mathrm{e}^{-\beta\mathcal{H}^{\mathsf{CW}}_N(\nu)}\right)}. \tag{4.2}$$

One way to pin down phase transition in the CW model is to study statistical properties of the mean magnetization

$$\bar\nu_N \;\triangleq\; \frac{1}{N}\sum_i\nu_i,$$

under μ^{β}_N. As it is well known, for $\beta\leq1$, the distribution of $\bar\nu_N$ is sharply concentrated around $\pm m^*$, where the spontaneous magnetization $m^*=m^*(\beta)$ equals to zero for $\beta\leq1$ and is positive (and hence

there are coexisting $\pm$ phases) for $\beta > 1$. This could be verified in two different ways, which correspond to two equality signs in (4.2): either directly through large deviation computations for Bernoulli random variables, or using the geometric FK representation as described in Subsection 2.3. In the latter case phase transition in the CW model is related to emergence of the giant component in the classical Erdős-Rényi random graph. Both methods are briefly recalled in Subsection 4.1

The main objective of this Section, however, is to explain that a very similar story happens with the quantum CW model in transverse field,

$$-\mathcal{H}_N^{\mathsf{CW}} \,=\, \frac{1}{N}\sum_{(i,j)} \hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}} \,+\, \lambda \sum_i \hat{\sigma}_i^{\mathsf{x}}.$$

In particular, there is a natural inclusion of (one parameter) Erdős-Rényi random graph models into a two-parameter family of space-time random graphs. In this way classical Erdős-Rényi critical point $\beta = 1$ is just the limiting point on the whole critical curve in the (β, λ) plane. It is somewhat amusing that, apparently, such quantum version of Erdős-Rényi random graphs was overlooked for a long time, and the corresponding critical curve was originally computed only in [15].

Contrary to what happens in the classical case, however, for the moment it is not clear how recover the critical curve for the quantum CW model in the transverse field from the critical curve for the quantum Erdős-Rényi random graph, although a conjecture has appeared in [14]. In principle, the quantum CW critical curve could be derived from the results of [17], where limiting states were classified for essentially all mean field type models. Alternatively, one can use infinite dimensional theory of large deviations, see [11] and references therein. In the concluding Subsection 4.3 we shall briefly report on recent results of [9]. As in [11] the approach relies on a partial Trotterization of the mean-field Hamiltonian under, however, a different choice of arrival operators associated to transversal field: Ours corresponds to the FK setup of Subsection 3.1. Such FK point of view leads to certain advantages and, as a result, we go beyond just computing the critical curve itself. In particular, we are able to derive sharp asymptotics of the spontaneous magnetization $m^*(\beta, \lambda)$ in the vicinity of the critical curve, and for (β, λ) away from the critical curve we are able to derive quadratic stability bounds for maximizers of the corresponding infinite dimensional mean-field variational problem.

4.1 Classical Case

The probability measure ν_N^β in (4.2) could be described in the following way: Let $\mathbb{Q}$ be the uniform (1/2) distribution on $\{\pm 1\}$ and let $\otimes\mathbb{Q}$ be the corresponding product measure on $\Omega_N = \{\pm 1\}^N$. Then,

$$\mu_N^\beta(\nu) \;=\; \frac{\otimes\mathbb{Q}\big(\mathrm{e}^{N\beta(\bar\nu_N)^2/2};\nu\big)}{\otimes\mathbb{Q}\left(\mathrm{e}^{N\beta(\bar\nu_N)^2/2}\right)}. \tag{4.3}$$

Then, elementary one-dimensional theory of large deviations implies that μ_N^β exponentially concentrates around

$$\Big\{\nu \,:\, \bar\nu_N \text{ is close to } \operatorname{argmax}\Big(\frac{\beta}{2}m^2 - I(m)\Big)\Big\},$$

where I is the large deviation rate function for $\bar\nu_N$ under $\otimes\mathbb{Q}$,

$$I(m) \;=\; \sup_h\{hm - \Lambda(h)\} \quad \text{and} \quad \Lambda(h) = \log\mathbb{Q}\big(\mathrm{e}^{h\nu}\big) = \log\frac{\mathrm{e}^h + \mathrm{e}^{-h}}{2}.$$

It is easy to see that I is strictly convex and differentiable on $(-1,1)$ with $I'(m) \to \pm\infty$ as $m \to \pm 1$. In particular, the supremum of $\beta m^2/2 - I(m)$ is actually attained inside $(-1,1)$ for any $\beta \in \mathbb{R}_+$. Furthermore, since $I(\cdot)/\beta$ is the convex conjugate of $\Lambda(\beta\cdot)/\beta$,

$$\operatorname{argmax}\left\{\frac{m^2}{2} - \frac{1}{\beta}I(m)\right\} \;=\; \operatorname{argmax}\left\{\frac{1}{\beta}\Lambda(\beta h) - \frac{h^2}{2}\right\}. \tag{4.4}$$

But $\Lambda(\beta\cdot)$ is the log-moment generating function of the $\pm\beta$ Bernoulli random variable. If we use $\mathbb{Q}_\beta$ for the corresponding distribution, then it is straightforward to check that the maximizers in (4.4) are of the form $\pm m^*(\beta)$, where $m^*(\beta) > 0$ iff,

$$1 \;<\; \frac{1}{\beta}\mathbb{V}\mathrm{ar}\;(\beta)\,(\nu) \;=\; \frac{\beta^2}{\beta}, \tag{4.5}$$

and we, thereby, recover the critical value $\beta = 1$ of the classical CW model.

Relation to Random Graphs. Let us go back to the definition of the classical FK measure in (2.31), and let us use the shorthand notation $\widetilde{\mathbb{P}}_{\beta,N}$ for the CW case at zero magnetic field, $J \equiv 1/N$ and $h = 0$. By the second equality in (4.2), the distribution μ_N^β can be constructed from $\widetilde{\mathbb{P}}_{\beta,N}$ as follows:

First sample arrival processes $\xi = \{\xi_{ij}\}$ from $\widetilde{\mathbb{P}}_{\beta,N}$. Two sites i and j (or, equivalently, two circles $\mathbb{S}^i_\beta$ and $\mathbb{S}^j_\beta$) are said to be connected in ξ if $\xi_{ij} \neq \emptyset$. Thus, any realization of ξ splits $\{1, \dots, N\}$ into maximal connected components. At the second step paint those connected components into ± 1 independently and with probability $1/2$ each. In fact we have just constructed a joint measure $\mathbb{M}_{\beta,N}(\mathrm{d}\xi, \nu)$ with marginals $\widetilde{\mathbb{P}}_{\beta,N}$ and μ^β_N.

In view of such two-step construction of μ^β_N, the critical point $\beta = 1$ and the value of the spontaneous magnetization $m^*(\beta)$ could be recovered now from the following facts about the FK measures $\widetilde{\mathbb{P}}_{\beta,N}$ on complete graph: With $\widetilde{\mathbb{P}}_{\beta,N}$-probabilities tending to 1, as N tends to ∞,

1. For $\beta < 1$ all connected components of ξ have sizes $\mathrm{O}(\log N)$ at most.
2. For $\beta > 1$, there is exactly one giant connected component of size $\sim m^*(\beta)N$, whereas the remaining connected components of ξ have sizes $\mathrm{O}(\log N)$ at most.

Above statements are similar to classical results on the emergence of giant component in random complete graphs. Indeed, by construction,

$$\widetilde{\mathbb{P}}_{\beta,N}(\mathrm{d}\xi) = \frac{2^{\#(\xi)}\mathbb{P}_{\beta,N}(\mathrm{d}\xi)}{\mathbb{P}_{\beta,N}\left(2^{\#(\xi)}\right)}, \tag{4.6}$$

where $\#(\xi)$ is the number of connected components of ξ (recall that since we take $h = 0$ there are no wired components as in the general formula (2.31)). We can think about $\mathbb{P}_{\beta,N}$ in terms of Erdős-Rényi random graph on $\{1, \dots, N\}$ where bonds between different sites i, j are placed independently and with probability $2\beta/N$ each. Indeed, $1 - \mathrm{e}^{-2\beta/N}$ is the probability that $\xi_{ij} \neq \emptyset$. Furthermore, as it was observed by Edwards and Sokal [13], the conditional ξ-marginal of

$$\mathbb{M}_{\beta,N}(\cdot|\nu_1 = 1, \dots, \nu_M = 1, \nu_{M+1} = -1, \dots, \nu_N = -1)$$

is exactly $\mathbb{P}_{\beta,M} \otimes \mathbb{P}_{\beta,N-M}$. Since $\max\{M, N-M\} \geq N/2$, the inequality $\beta_c \leq 1$ for the critical FK value of β is immediately implied by classical Erdős-Rényi results, see e.g. [6]: Let $\{1, \dots, K\}$ be the complete graph of K sites. Assume that an (un-oriented) edge (i, j) is open with probability ϵ/K independently from all other edges. Then $\epsilon_c = 1$ is the threshold for the emergence of the giant component. Moreover,

in the case of $\epsilon > 1$ the density $\rho(\epsilon)$ of the giant component is asymptotically close to the positive solution of

$$1 - \rho = \mathrm{e}^{-\epsilon\rho}. \tag{4.7}$$

In over case, $K = \max\{M, N-M\} \geq N/2$, and hence $2\beta/N > 1/K$ whenever $\beta > 1$.

The reverse inequality $\beta_c \geq 1$ is not much harder: Assume that $\beta < 1$. Without loss of generality we can consider only the case when the total number of $+$spins $M \leq N/2$. Then, under $\mathbb{P}_{\beta,M}$ all the connected components of $\{1, \ldots, M\}$ are small. A-priori, a giant connected component still could appear under $\mathbb{P}_{\beta,N-M}$. Let ρ be the density of this component. Then $(1-\rho)(N-M)$ of the remaining $-$spins live on small components of sizes $\mathrm{O}(\log N)$ at most. Since in the original coupled measure $\mathbb{M}_{\beta,N}$ all the small connected components were coloured independently, we infer that $M \sim (1-\rho)(N-M)$. Accordingly, $K \stackrel{\Delta}{=} N - M \sim N/(2-\rho)$ and hence $2\beta/N \sim \epsilon/K$ with $\epsilon = 2\beta/(2-\rho)$. Thus, by (4.7), the relative density ρ should satisfy

$$1 - \rho = \mathrm{e}^{-2\beta\rho/(2-\rho)}.$$

But the latter equation does not have a positive solution. Indeed, set $\theta = \rho/(2-\rho)$ or $\rho = (1-\theta)/(1+\theta)$. Then θ is positive as soon as ρ is positive, and

$$\frac{1-\theta}{1+\theta} = \mathrm{e}^{-2\beta\theta}$$

Taking logs and expanding,

$$2\theta + \frac{2}{3}\theta^3 + \ldots = 2\beta\theta,$$

which is impossible unless $\theta =$ or $\beta > 1$.

A general class of FK models on complete graphs is examined in [7].

4.2 Curie-Weiss Model in Transverse Field and Quantum Random Graphs

Quantum Curie-Weiss Hamiltonian in transverse field $\lambda \geq 0$ is given by,

$$-\mathcal{H}_N^{\mathsf{CW}} = \frac{1}{N}\sum_{(i,j)} \hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}} + \lambda \sum_i \hat{\sigma}_i^{\mathsf{x}}.$$

Following the approach of Subsection 3.1 we associate to $\mathcal{H}_N^{\mathsf{CW}}$ the following family ξ of independent Poisson processes of arrivals on the

circle $\mathbb{S}_\beta$: For each un-oriented couple (i,j) operators $\left(\mathrm{I}+\hat{\sigma}_i^{\mathsf{z}}\hat{\sigma}_j^{\mathsf{z}}\right)$ arrive with intensity $2/N$, whereas operators $(\mathrm{I}+\hat{\sigma}_i^{\mathsf{x}})$ arrive with intensity λ for every $i=1,\ldots,N$. Connected components of $\{1,\ldots,N\}\times\mathbb{S}_\beta$ induced by ξ are defined precisely as in Subsection 3.1. Recall that each such connected component $\mathcal{C}$ is represented as a union,

$$\mathcal{C}\,=\,\bigcup_l\{i_l\times\ I_l\}\,,$$

of disjoint space-time intervals. The size of $\mathcal{C}$ could be measured in several ways: For example we can compute number of different spatial coordinates (out of $\{1,\ldots,N\}$) which contribute to $\mathcal{C}$. The most natural definition of the size, however, is

$$|\mathcal{C}|\,=\,\sum_l|I_l|, \tag{4.8}$$

that is the total length of all time intervals of $\mathcal{C}$.

Since we consider the case of zero z-field, all connected components of ξ are free. Consequently, the FK modification $\widetilde{\mathbb{P}}^\lambda_{\beta,N}$ of the reference product Poisson measure $\mathbb{P}^\lambda_{\beta,N}$ is given by

$$\widetilde{\mathbb{P}}^\lambda_{\beta,N}\,(\mathrm{d}\xi)\,=\,\frac{2^{\#(\xi)}\mathbb{P}^\lambda_{\beta,N}\,(\mathrm{d}\xi)}{\mathbb{P}^\lambda_{\beta,N}\left(2^{\#(\xi)}\right)}, \tag{4.9}$$

In view of (3.7) it is suggestive to try to study the question of phase co-existence in terms of emergence of giant components under $\widetilde{\mathbb{P}}^\lambda_{\beta,N}$. Note that in a genuine quantum case of $\lambda>0$, this is a non-trivial question even in the ground state limit when $\beta\to\infty$. In fact, instead of one critical value of β one should face here a whole critical curve in the (λ,β) positive quarter plane. For the moment we do not know how to derive this curve via direct analysis of random space-time graphs induced by the family of quantum FK measures (4.9). This, however, is a meaningful question even for the reference family of measures $\mathbb{P}^\lambda_{\beta,N}$.

Quantum Random Graphs As it is apparent from a comparison between (4.9) and (4.6) the measures $\mathbb{P}^\lambda_{\beta,N}$ play the same role for the quantum Curie-Weiss model in transverse field as Erdős-Rényi random graphs $\mathbb{P}_{\beta,N}$ play for the classical CW model. Accordingly, we shall refer to the collection of independent Poisson processes of holes and links induced by $\mathbb{P}^\lambda_{\beta,N}$ as to quantum random graphs. In order to be compatible with the usual random graph notation let us modify the arrival rates under $\mathbb{P}^\lambda_{\beta,N}$ in the following way: The holes still arrive with intensity λ,

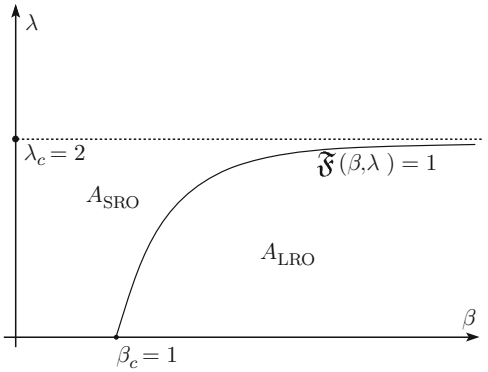


Fig. 5. Decomposition of the (β, λ) quarter plane into the short range and long range regions.

however the links between an unordered pair of sites $i \neq j$ arrive now with intensity $1/N$. In this way $\beta_c = 1$ is the classical critical value which corresponds to $\lambda = 0$. The main result of [15] asserts that the full critical curve for the family of quantum random graphs is implicitly given by,

$$\mathfrak{F}(\beta, \lambda) \stackrel{\Delta}{=} \frac{2}{\lambda}\left(1 - \mathrm{e}^{-\lambda\beta}\right) - \beta \mathrm{e}^{-\lambda\beta} = 1. \tag{4.10}$$

The curve is depicted on Figure 5. Note that the classical critical value $\beta_c = 1$ is just the end-point of the curve on β-axis. Notice also that the critical value of λ in the ground state model $\beta = \infty$ equals to $\lambda_c = 2$. Let us be more specific about the nature of phase transition for quantum random graphs: The critical curve (4.10) splits the positive quarter-plane into

$$A_{\mathrm{LRO}} \stackrel{\Delta}{=} \{(\beta, \lambda) : \mathfrak{F}(\beta, \lambda) > 1\} \text{ and } A_{\mathrm{SRO}} \stackrel{\Delta}{=} \{(\beta, \lambda) : \mathfrak{F}(\beta, \lambda) < 1\},$$

where LRO (respectively SRO) stands for long (respectively short) range order. Here is a justification for such a terminology: By definition, two points $(i, t), (j, s)$ are connected in ξ, if the intervals containing these points belong to the same connected component $\mathcal{C}$ in the ξ-induced decomposition of $\{1, \ldots, N\} \times \mathbb{S}_\beta$. We shall denote the latter event as $\{(i, t) \longleftrightarrow (j, s)\}$. Then,

1. If $(\beta, \lambda) \in A_{\mathrm{SRO}}$, then

$$\mathbb{P}^{\lambda}_{\beta,N}\left((i, t) \longleftrightarrow (j, s)\right) = O\Big(\frac{\log N}{N}\Big) \tag{4.11}$$

uniformly in $t, s \in \mathbb{S}_\beta$ and $i \neq j$.

2. On the other hand, if $\beta < \infty$ and $(\beta, \lambda) \in A_{\mathrm{LRO}}$, then there exists $\rho = \rho(\beta, \lambda) \in (0, 1)$, such that

$$\mathbb{P}^{\lambda}_{\beta,N}\left((i,t) \longleftrightarrow (j,s)\right) \,=\, \rho(\beta,\lambda)^2\left(1 + o(1)\right), \tag{4.12}$$

also uniformly in $t, s \in \mathbb{S}_\beta$ and $i \neq j$.

As in the classical Erdős-Rényi case the short/long range order transition for $\beta < \infty$ is related to an emergence of a unique giant connected component. In fact, the number $\rho(\beta, \lambda)$ in (4.12) is precisely the limiting space-time density of the latter. More precisely, let us use (4.8) to measure sizes of random connected components of $\{1, \dots, N\} \times \mathbb{S}_\beta$. Let $\mathcal{M}$ and $\mathcal{M}^{\mathrm{next}}$ be the largest and the next to the largest sizes of these connected components (of course, these definitions make sense only for $\beta < \infty$). Then,

1. If $(\beta, \lambda) \in A_{\mathrm{SRO}}$, then for every $\kappa > 0$ there exists $c = c(\beta, \lambda, \kappa) < \infty$, such that

$$\mathbb{P}^{\lambda}_{\beta,N}\left(\left|\mathcal{C}\left((i,t)\right)\right| > c\log N\right) \,=\, o\Big(\frac{1}{N^{\kappa}}\Big), \tag{4.13}$$

where $\mathcal{C}((i,t))$ is the connected component containing (i,t). Clearly, the distribution of $\left|\mathcal{C}\left((i,t)\right)\right|$ is the same for all $i \in \{1, \dots, N\}$ and $t \in \mathbb{S}_\beta$ (by definition $\mathbb{S}_\infty = \mathbb{R}$). Furthermore, if $\beta < \infty$, then

$$\mathbb{P}^{\lambda}_{\beta,N}\left(\mathcal{M} > c\log N\right) \,=\, o\Big(\frac{1}{N^{\kappa-1}}\Big) \tag{4.14}$$

2. If, however, $\beta < \infty$ and $(\beta, \lambda) \in A_{\mathrm{LRO}}$ then there exists a sequence of positive numbers $\epsilon_N(\beta, \lambda) \to 0$ such that,

$$\mathbb{P}^{\lambda}_{\beta,N}\left(\left|\frac{|\mathcal{C}((i,t))|}{N\beta} - \rho\right| < \epsilon_N\right) \,=\, \rho(\beta,\lambda)(1 - o(1)), \tag{4.15}$$

where $\rho(\beta, \lambda)$ is the same probability as in (4.12). Furthermore, in the $\beta < \infty$ case, there exists a constant $c = c(\beta, \lambda) < \infty$ such that

$$\mathbb{P}^{\lambda}_{\beta,N}\left(\mathcal{E}(\rho, \epsilon_N, c)\right) = 1 - o(1), \tag{4.16}$$

where the event $\mathcal{E}(\rho, \epsilon_N, c)$ is defined via

$$\mathcal{E}(\rho, \epsilon_N, c) \,=\, \left\{\left|\frac{\mathcal{M}}{\beta N} - \rho\right| < \epsilon_N\right\} \cap \left\{\mathcal{M}^{\mathrm{next}} < c\log N\right\}. \tag{4.17}$$

The original proof of the above results appeared in [15]. Afterwards, the statements related to the $\beta < \infty$ case were re-proven using somewhat different methods in [16].

We finish this Subsection by indicating how the expression (4.10) comes into play. As in the classical case one couples a construction of a single connected component with a Galton-Watson process. In the quantum case descendant of a point $(i,t) \in \{1,\dots,N\}$ are generated in the following fashion:

1. First generate a random interval $I \subseteq \mathbb{S}_\beta$ around (i,t),so that the end-points of I would imitate two successive holes. Since the holes arrive with intensity λ the length $|I|$ should be distributed as $\min\{\Gamma(2,\lambda),\beta\}$.
2. Given a realization of $I \ni t$, the number of all links to i which arrive during I is distributed Poisson($\frac{N-1}{N}|I|$). In the Galton-Watson approximation we take it to be exactly Poisson($|I|$).

Accordingly, if we denote the number of descendants in the Galton-Watson approximation by X, then $\mathbb{E}(X|I) = |I|$. Let $V \sim \Gamma(2,\lambda)$ Then,

$$\mathbb{E}\left(|I|\right) \,=\, \mathbb{E}\left(V; V < \beta\right) + \beta\mathbb{P}\left(V \geq \beta\right),$$

Now,

$$\mathbb{P}\left(V \geq \beta\right) \,=\, \int_\beta^\infty \lambda^2 t \mathrm{e}^{-\lambda t}\mathrm{d}t \,=\, (\lambda\beta+1)\mathrm{e}^{-\lambda\beta}.$$

In the same fashion,

$$\mathbb{E}\left(V; V \leq \beta\right) \,=\, \frac{2}{\lambda}\Big(1-\mathrm{e}^{-\lambda\beta}\Big) \,-\, \left(\beta^2\lambda+2\beta\right)\mathrm{e}^{-\lambda\beta}.$$

Consequently,

$$\mathbb{E}\left(|I|\right) \,=\, \frac{2}{\lambda}\Big(1-\mathrm{e}^{-\lambda\beta}\Big) \,-\, \beta\mathrm{e}^{-\lambda\beta},$$

which is precisely the expression in (4.10).

4.3 Critical Curve for Quantum Curie-Weiss Model via Large Deviations

Large deviation representation of the CW model in transverse field is obtained via partial linearization in the Lie-Trotter product formula (or partial Poissonization of the CW Hamiltonian). Namely,

$$\frac{\mathrm{e}^{-\beta\mathcal{H}_N^{\mathsf{CW}}}}{\mathrm{e}^{\lambda N}} = \lim_{\Delta\to 0}\left(\prod_{(i,j)} \mathrm{e}^{\frac{\Delta}{N}\hat{\sigma}_i^z\hat{\sigma}_j^z}\prod_i\left\{(1-\Delta\lambda)\mathrm{I}+\Delta\lambda(\hat{\sigma}_i^{\mathsf{x}}+\mathrm{I})\right\}\right)^{\beta/\Delta}. \quad (4.18)$$

Note that matrices $\mathrm{e}^{\frac{\Delta}{N}\hat{\sigma}_i^z\hat{\sigma}_j^z}$ are diagonal in the z-basis,

$$\langle\Psi_\nu|\mathrm{e}^{\frac{\Delta}{N}\hat{\sigma}_i^z\hat{\sigma}_j^z}|\Psi_{\nu'}\rangle \;=\; \mathrm{e}^{\frac{\Delta}{N}\nu_i\nu_j}. \tag{4.19}$$

Let $\mathbb{P}_\beta^\lambda$ be the distribution of the Poisson point process (of holes) on the circle $\mathbb{S}_\beta$ with arrival intensity λ. We shall use $\otimes\mathbb{P}_\beta^\lambda$ for the product distribution of N independent copies $\xi = (\xi_1, \dots, \xi_N)$. Given a realization of ξ let us say that a classical piece-wise constant trajectory $\nu : \mathbb{S}^\beta \mapsto \{\pm 1\}^N$ is compatible with ξ; $\nu \sim \xi$, if for every $i = 1, \dots, N$ jumps of $\nu_i(\cdot)$ occur only at arrival times of ξ_i. Passing to the limit in (4.18) we, in view of (4.19), infer

$$\frac{\mathsf{Tr}\big(\mathrm{e}^{-\beta\mathcal{H}_N^{\mathsf{CW}}}\big)}{\mathrm{e}^{\lambda N}} = \int \otimes\mathbb{P}_\beta^\lambda(\mathrm{d}\xi) \sum_{\nu\sim\xi} \exp\Big\{\int_0^\beta \frac{1}{N}\sum_{(i,j)} \nu_i(t)\nu_j(t)\mathrm{d}t\Big\}. \tag{4.20}$$

For every i let $\#(\xi_i)$ be the number of connected components of $\mathbb{S}_\beta \setminus \xi_i$. Evidently, the number of all compatible $\nu \sim \xi$ equals to $2^{\sum_i \#(\xi_i)}$. Define

$$\widetilde{\mathbb{P}}_\beta^\lambda\,(\mathrm{d}\xi) \;=\; \frac{2^{\#(\xi)}\mathbb{P}_\beta^\lambda\,(\mathrm{d}\xi)}{\mathbb{P}_\beta^\lambda\,\big(2^{\#(\xi)}\big)}$$

This is just the one-circle FK measure. Consider probability distribution $\mathbb{Q}_\beta^\lambda$ on piece-wise constant classical one-circle spin trajectories $\nu : \mathbb{S}_\beta \mapsto \{\pm 1\}$ which is generated by the following two step procedure: First sample ξ from $\widetilde{\mathbb{P}}_\beta^\lambda$, and then paint connected components of $\mathbb{S}_\beta \setminus \xi$ into ± 1, independently and with probability $1/2$ each. Let $\otimes\mathbb{Q}_\beta^\lambda$ be the corresponding product measure. It is straightforward to check that the righthand side of (4.20) equals to

$$\Big[\widetilde{\mathbb{P}}_\beta^\lambda\big(\mathrm{e}^{\#(\xi)}\big)\Big]^N \otimes \mathbb{Q}_\beta^\lambda\Big(\exp\Big\{\int_0^\beta \frac{1}{N}\sum_{(i,j)} \nu_i(t)\nu_j(t)\mathrm{d}t\Big\}\Big).$$

Consequently, an analysis of phase diagram of the CW model in transverse filed boils down to an investigation of asymptotic properties for weighted measures

$$\otimes\widetilde{\mathbb{Q}}_\beta^\lambda(\mathrm{d}\nu) \;\stackrel{\Delta}{=}\; \frac{\otimes\mathbb{Q}_\beta^\lambda\left(\exp\left\{\frac{N}{2}\int_0^\beta \left(\bar{\nu}_N(t)\right)^2\mathrm{d}t\right\}\,;\,\mathrm{d}\nu\right)}{\otimes\mathbb{Q}_\beta^\lambda\left(\exp\left\{\frac{N}{2}\int_0^\beta \left(\bar{\nu}_N(t)\right)^2\mathrm{d}t\right\}\right)}, \tag{4.21}$$

where,

$$\bar{\nu}_N(t) \;=\; \frac{1}{N}\sum_i \nu_i(t).$$

This problem belongs to the realm of theory of large deviations. Formally, the measures (4.21) are asymptotically concentrated around solutions of

$$\sup_m\Big\{\frac{1}{2}\int_0^\beta m^2(t)\mathrm{d}t \;-\; I(m)\Big\} \;\stackrel{\Delta}{=}\; \sup_m \mathfrak{G}(m), \tag{4.22}$$

where I is the large deviation rate function for the average $\bar{\nu}_N$ under the product measures $\otimes\mathbb{Q}_\beta^\lambda$. If we formulate the large deviation principle in $\mathbb{L}_2(\mathbb{S}_\beta)$, then, using $(\cdot,\cdot)_\beta$ for the corresponding scalar product,

$$I(m) \;=\; \sup_h\{(h,m)_\beta - \Lambda(h)\} \quad \text{where} \quad \Lambda(h) \;=\; \log\mathbb{Q}_\beta^\lambda\big(\mathrm{e}^{(h,\nu)_\beta}\big). \tag{4.23}$$

A detailed analysis of the variational problem (4.22) and of the weighted measures $\widetilde{\mathbb{Q}}_{\beta,N}^\lambda$ will appear in the forthcoming [9]. Here we shall try to give a brief sketch of the results and techniques, in particular, we shall explain how the critical curve of the CW model in the transverse field could be read from (4.22).

The critical curve is implicitly given by

$$\mathfrak{f}(\lambda,\beta) \;\stackrel{\Delta}{=}\; \frac{1}{\beta}\mathbb{V}\mathrm{ar}_\lambda\,(\beta)\,((\nu,\mathbb{1})_\beta) \;=\; \frac{1}{\lambda}\tanh(\lambda\beta) \;=\; 1, \tag{4.24}$$

where $\mathbb{V}\mathrm{ar}_\lambda\,(\beta)$ is the variance under the one-circle spin measure $\mathbb{Q}_\beta^\lambda$. As we show in [9], the variational problem (4.22) has constant maximizers $\pm m^*(\lambda,\beta)$, where the spontaneous z-magnetization m^* satisfies:

1. If $\mathfrak{f}(\lambda,\beta)\le 1$, then $m^*=0$.
2. If $\mathfrak{f}(\lambda,\beta)>1$, then $m^*>0$, and, consequently there are two distinct solutions to (4.22).

Furthermore, away from the critical curve the solutions $\pm m^*\mathbb{1}$ are stable in the following sense: There exists $c=c(\lambda,\beta)>0$ and a strictly convex symmetric function U with $U(r)\sim r\log r$ growth at infinity such that

$$\mathfrak{G}(\pm m^*\mathbb{1})-\mathfrak{G}(m)\ge c\min\big\{\|m-m^*\mathbb{1}\|_\beta^2,\|m+m^*\mathbb{1}\|_\beta^2\big\}+\int_0^\beta U(m'(t))\mathrm{d}t. \tag{4.25}$$

The second term above is important in the super-critical regime ($\mathfrak{f}(\lambda,\beta)>1$) since it rules out trajectories of $\bar{\nu}_N(\cdot)$ with rapid transitions between the optimal values $\pm m^*$.

Properties of One-circle Spin Measures

The following properties of $\mathbb{Q}^{\lambda}_{\beta}$ are crucial for the analysis of (4.22):

1. $\mathbb{Q}^{\lambda}_{\beta}$ possesses the FKG property.
2. $\mathbb{Q}^{\lambda}_{\beta}$ satisfies the following qualitative version of the GHS inequality: Given $h \in \mathbb{R}_+$ define the tilted measure

$$\mathbb{Q}^{\lambda,h}_{\beta}(\mathrm{d}\nu) \;=\; \frac{\mathbb{Q}^{\lambda}_{\beta}\left(\mathrm{e}^{h(\nu,\mathbb{1})_{\beta}}\,;\,\mathrm{d}\nu\right)}{\mathbb{Q}^{\lambda}_{\beta}\left(\mathrm{e}^{h(\nu,\mathbb{1})_{\beta}}\right)}.$$

   Then, there exists $c_1 = c_1(\lambda,\beta) > 0$, such that

$$\frac{\mathrm{d}}{\mathrm{d}h}\mathbb{V}\mathrm{ar}_{\lambda,h}(\beta)\left((\nu,\mathbb{1})_{\beta}\right) \;\le\; -c_1 h\mathrm{e}^{-2\beta h}. \tag{4.26}$$

3. $\mathbb{Q}^{\lambda}_{\beta}$ is reflection positive: Let $n \in \mathbb{N}$, $0 < t_1 < \cdots < t_n < \beta/2$ and let $f : \{\pm 1\}^n \to \mathbb{C}$. Set $s_k = \beta - t_k$. Then,

$$\mathbb{Q}^{\lambda}_{\beta}\left(f(\nu_{t_1},\ldots,\nu_{t_n})\bar{f}(\nu_{s_1},\ldots,\nu_{s_n})\right) \;\ge\; 0. \tag{4.27}$$

Properties 1. and 3. are more or less immediate since $\mathbb{Q}^{\lambda}_{\beta}$ could be viewed in terms of an approximation by ferromagnetic nearest neighbour one-dimensional Ising models. Namely, let us approximate ξ by Bernoulli point process of arrivals ξ^{Δ}, exactly as in (2.21). Modify Bernoulli weights by $2^{\#(\xi^{\Delta})}$ and paint connected components of $\mathbb{S}\setminus\xi^{\Delta}$ into ± 1, independently and with probability $1/2$ each. Then, the resulting measure $\mathbb{Q}^{\lambda}_{\beta,\Delta}$ approximates $\mathbb{Q}^{\lambda}_{\beta}$. Of course $\mathbb{Q}^{\lambda}_{\beta,\Delta}$ charges only trajectories ν which jump at times $j\Delta$. For such trajectories,

$$\mathbb{Q}^{\lambda}_{\beta,\Delta}(\nu) \;\sim\; \prod_{i=0}^{\beta/\Delta-1}\left(\delta_{\{\nu(i\Delta)=\nu((i+1)\Delta)\}} + \Delta\lambda\delta_{\{\nu(i\Delta)=\nu(i+1)\Delta\}}\right).$$

Set $J = J(\Delta,\lambda) = -\log\sqrt{\Delta\lambda}$. Since

$$\delta_{\{\nu(i\Delta)=\nu((i+1)\Delta)\}} + \Delta\lambda\delta_{\{\nu(i\Delta)=\nu((i+1)\Delta)\}} \;=\; \frac{\mathrm{e}^{J\nu(i\Delta)\nu((i+1)\Delta)}}{\mathrm{e}^{J}},$$

we recognize $\mathbb{Q}^{\lambda}_{\beta,\Delta}$ as a scaling of the nearest neighbour Ising model on discrete circle $\mathbb{S}_{\beta}/\Delta$ at unit temperature and with interaction strength $J(\Delta,\lambda)$.

Inequality (4.26) is proved in [9] using the same approximation (by 1D Ising models) with an additional care being paid to limits of random current representation of third semi-invariants (based on [2]).

Dual Variational Problem

In order to explain the implications of the properties of $\mathbb{Q}_\beta^\lambda$ listed above, it is convenient to consider the dual variational problem,

$$\sup_h \Big\{ \Lambda(h) \, - \, \frac{1}{2} \int_0^\beta h^2(t) \mathrm{d}t \Big\} \stackrel{\Delta}{=} \sup_h \mathfrak{G}^*(h). \tag{4.28}$$

Any solution $\tilde{h}$ of (4.28) is also a solution to (4.22). This is a general fact from convex analysis: Let F and G be two proper lower-semicontinuous convex functionals (on say $\mathbb{L}_2(\mathbb{S}_\beta)$) and let F^* and G^* be their convex conjugates. Assume that

$$F^*(\tilde{h}) - G^*(\tilde{h}) \, = \, \max_h \{F^*(h) - G^*(h)\} \, ,$$

and assume that both F^* and G^* are Gateaux differentiable (in fact sub-differentiability would be enough) at $\tilde{h}$. Let $\tilde{m} = \nabla F^*(\tilde{h}) = \nabla G^*(\tilde{h})$. Then,

$$F^*(\tilde{h}) - G^*(\tilde{h}) = G(\tilde{m}) - F(\tilde{m}).$$

Consequently, for each couple of functions m and h,

$$\{(m,h)_\beta - G^*(h)\} - \{(m,h)_\beta - F^*(h)\} \leq G(\tilde{m}) - F(\tilde{m}).$$

It follows that for every m, $G(m) - F(m) \leq G(\tilde{m}) - F(\tilde{m})$. Furthermore, assume that we can quantify stability property of the dual variational problem in the following way: There exists a non-negative functional D, such that $D = 0$ only on the solutions of the dual problem, and for any function h,

$$F^*(h) - G^*(h) + D(h) \, \leq \, F^*(\tilde{h}) - G^*(\tilde{h}). \tag{4.29}$$

Then such stability bound is transferable to the direct problem: Assume that $h = \nabla G(m)$. Then,

$$G(m) - F(m) + D(h) + \{F(m) + F^*(h) - (m,h)_\beta\} \, \leq \, G(\tilde{m}) - F(\tilde{m}). \tag{4.30}$$

In particular, $G(m) - F(m) < G(\tilde{m}) - F(\tilde{m})$, whenever $\nabla G(m)$ *is not* a solution of the dual problem or whenever $h \notin \partial F(m)$.

Let us now go back to (4.22) and (4.28). In the above notation: $F(m) = I(m)$ and $G(m) = \|m\|_\beta^2/2$. Accordingly, $F^*(h) = \Lambda(h)$ and $G^*(h) = \|h\|_\beta^2/2$. In particular, G, G^* and F^* are everywhere Gateaux differentiable. Of course, $\nabla G(m) = m$. Consequently, once we derive a

stability bound of the type (4.29) for the dual problem, we immediately recover a stability bound

$$\frac{1}{2}\int_0^\beta m^2(t)\mathrm{d}t - I(m) + D(m) + \left\{I(m) + \Lambda(m) - \|m\|_\beta^2\right\} \le \mathfrak{G}(\tilde{m}) \tag{4.31}$$

for the original problem (4.22). In particular, any solution of (4.22) is a solution of (4.28).

We, therefore, proceed to study the dual variational problem (4.28).

Reduction to a One-dimensional Problem

Reflection positivity property (4.27) implies that for any $h \in \mathbb{L}_2(\mathbb{S}_\beta)$,

$$\Lambda(h) \le \frac{1}{\beta}\int_0^\beta \Lambda\left(h(t)\mathbb{1}\right)\mathrm{d}t. \tag{4.32}$$

Note that (4.32) has been originally proved in a somewhat more general context in [11]. As a result,

$$\mathfrak{G}^*(h) \le \int_0^\beta \Big\{\frac{1}{\beta}\Lambda\left(h(t)\mathbb{1}\right) - \frac{1}{2}h^2(t)\Big\}\mathrm{d}t \le \beta \sup_{h\in\mathbb{R}}\Big\{\frac{1}{\beta}\Lambda\left(h\mathbb{1}\right) - \frac{1}{2}h^2\Big\}.$$

We claim that the maximizers of the one-dimensional variational problem

$$\max_{h\in\mathbb{R}}\Big\{\frac{1}{\beta}\Lambda\left(h\mathbb{1}\right) - \frac{1}{2}h^2\Big\}, \tag{4.33}$$

are of the form $\pm h^*$, where $h^* > 0$ if and only if $\mathfrak{f}(\lambda,\beta) > 1$.

The critical curve (4.24). Compute,

$$\frac{\mathrm{d}}{\mathrm{d}h}\Big\{\frac{1}{\beta}\Lambda(h\mathbb{1}) - \frac{1}{2}h^2\Big\} = \frac{1}{\beta}\mathbb{Q}_\beta^{\lambda,h}\left((\nu,1)_\beta\right) - h.$$

The latter expression is evidently negative for h large enough, hence the maximum in (4.33) is attained at a critical point. Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}h}\mathbb{Q}_\beta^{\lambda,h}\left((\nu,1)_\beta\right) = \mathbb{V}\mathrm{ar}_\lambda\left(\beta\right)\left((\nu,1)_\beta\right).$$

Since by symmetry at $h = 0$ the expectation $\mathbb{Q}_\beta^{\lambda}\left((\nu,1)_\beta\right) = 0$, and since by (4.26) the function $h \to \mathbb{Q}_\beta^{\lambda,h}\left((\nu,1)_\beta\right)$ is strictly concave on $[0,\infty)$, we infer that:

Either $\mathbb{V}\mathrm{ar}_\lambda\left(\beta\right)\left((\nu,1)_\beta\right) \le \beta$, and then $h = 0$ is the only critical point of the function in (4.33). Or, $\mathbb{V}\mathrm{ar}_\lambda\left(\beta\right)\left((\nu,1)_\beta\right) > \beta$, and then

there are exactly three critical points; 0 and $\pm h^*$, the latter inevitably being the global maxima.

Stability of the one-dimensional problem. We claim, furthermore, that whenever (λ, β) is away from the critical curve, the problem (4.33) is stable,

$$\left\{\frac{1}{\beta}\Lambda\left(h\mathbb{1}\right) - \frac{1}{2}h^2\right\} + d(h) \;\leq\; \frac{1}{\beta}\Lambda\left(\pm h^*\mathbb{1}\right) - \frac{1}{2}(h^*)^2, \tag{4.34}$$

where d satisfies the following bound: There exists $c_1 = c_1(\lambda, \beta) > 0$, such that,

$$d(h) \geq c_2 \mathrm{e}^{-2\beta|h|} \min\left\{(h - h^*)^2, (h + h^*)^2\right\}. \tag{4.35}$$

Proof of (4.34). Follows from (4.26).

Stability of the original variational problem. It follows that the dual variational problem (4.28) (recall that in our case $F^*(\cdot) = \Lambda(\cdot)$ and $G^*(\cdot) = 1/2\|\cdot\|_\beta^2$) satisfies (4.29) with

$$D(h) \;=\; \frac{1}{\beta}\int_0^\beta d(h(t))\mathrm{d}t.$$

Of course, the bound (4.34) could be improved for large values of $|h|$, however since we are primarily interested in transferring stability to the direct variational problem (4.22), the values of $|h| > 1$ are, in view of (4.31), irrelevant. In particular $D(m)$ clearly dominates (with $h^* = m^*$ and c chosen appropriately small) the first term on the right hand side of (4.25).

The second term $\int_0^\beta U(m'(t))\mathrm{d}t$ on the right hand side of (4.25) is related to a more careful analysis of $\left\{I(m) + \Lambda(m) - \|m\|_\beta^2\right\}$ term in (4.31), which is unfortunately beyond the scope of these lectures. We, therefore, refer the reader to [9].

Behaviour Near the Critical Curve

The GHS-type bound (4.26) implies that the 4-th semi-invariant

$$-\mathsf{s}_4(\lambda, \beta) \;\triangleq\; \frac{\mathrm{d}^4\Lambda(h\mathbb{1})}{\mathrm{d}h^4}\Big|_{h=0},$$

is locally uniformly negative. Let $\mathfrak{f}(\lambda, \beta) > 1$ and assume that (λ, β) is close to the critical curve, in particular that $h^*(\lambda, \beta)$ is small. Then,

$$h^*(\lambda,\beta) = \frac{1}{\beta}\int_0^{h^*} \mathbb{V}\mathrm{ar}_{\lambda,\tau}\,(\beta)\,((\nu,\mathbb{1}))\,\mathrm{d}\tau$$
$$= h^*\mathfrak{f}(\lambda,\beta) - \frac{\mathsf{s}_4(\lambda,\beta)(h^*)^3}{6\beta}\,(1+\mathrm{O}(h^*))\,. \quad (4.36)$$

It follows that in the vicinity of the critical curve spontaneous magnetization $m^*(\lambda,\beta) = h^*(\lambda,\beta)$ scales like

$$\frac{m^*(\lambda,\beta)}{\sqrt{6\beta(\mathfrak{f}(\lambda,\beta)-1)/\mathsf{s}_4(\lambda,\beta)}} \;=\; 1+o(1).$$

References

1. Aizenman, M.: Geometric analysis of φ^4 fields and Ising models. I, II. Comm. Math. Phys. **86**, 1, 1–48 (1982).
2. Aizenman, M., Fernández, R.: On the critical behavior of the magnetization in high-dimensional Ising models. J. Stat. Phys., **44**, 3/4, 393–454 (1986).
3. Aizenman, M., Klein, A., Newman, C.M.: Percolation methods for disordered quantum Ising models, Mathematics, Physics, Biology..., R. Kotecký, ed., 1–24, World Scientific, Singapore (1993).
4. Aizenman, M., Nachtergaele, B.: Geometric aspects of quantum spin states. Comm. Math. Phys., **164**, 1, 17–63 (1994).
5. Björnberg, J.E., Grimmett, G: The phase transition of the quantum Ising model is sharp. http://xxx.lanl.gov/pdf/0901.0328
6. Bollobás, B.: Random Graphs, London: Academic Press (1985).
7. Bollobás, B., Grimmett, G., Janson, S.: The random cluster model on the complete graph. Prob. Theory Rel. Fields **104**, 3, 283–317 (1996).
8. Campanino, M., Klein, A., Perez, J.F.: Localization in the ground state of the Ising model with a random transverse field. Comm. Math. Phys. **135**, 3, 499–515 (1991).
9. Chayes, L., Crawford, N, Ioffe, D, Levit, A.: the phase diagram of the quantum Curie-Weiss model, J. stat. phys. **133**, 1, 131–149 (2008).
10. Crawford, N., Ioffe, D.: Random current representation for transverse field Ising models, http://xxx.lanl.gov/pdf/0812.4834
11. Dorlas, T.V.: Probabilistic derivation of a noncommutative version of Varadhan's theorem. Preprint DIAS-STP-02-5 (2002).
12. Duneau, M., Iagolnitzer, D., Souillard, B.: Strong cluster properties for classical systems with finite range interaction. Comm. Math. Phys. **35**, 307–320 (1974).
13. Edwards, R.G., Sokal, A.D.: Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm. Phys. Rev. D, **38**, 2009-2012 (1988).

14. Grimmett, G.: Space-time percolation, Preprint (2007).
15. Ioffe, D., Levit, A.: Long range order and giant components of quantum random graphs. Mark. Proc. and Rel. Fields **13**, 3, 469–492. (2007)
16. Janson, S.: On a random graph related to quantum theory. Combin. Probab. Comput. **16**, 757–766 (2007).
17. Fannes, M., Spohn, H., Verbeure, A.: Equilibrium states for mean field models. J. Math. Phys., **21**, 2, 355–358 (1980).
18. Nachtergaele, B.: Quasi-state decompositions for quantum spin systems. Probability theory and mathematical statistics (Vilnius, 1993), 565–590, TEV, Vilnius (1994).
19. Reed, M., Simon, B.: Methods of modern mathematical physics. I. Functional analysis. Second edition. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York (1980).
20. Ueltschi, D.: Geometric and probabilistic aspects of boson lattice models. In and out of equilibrium: Physics with a probability flavor, Progr. Probab. **51**, 363–391, Birkhäuser (2002).

Localization Transition in Disordered Pinning Models

Fabio Lucio Toninelli

Laboratoire de Physique, UMR-CNRS 5672, ENS Lyon,
46 Allée d'Italie, 69364 Lyon Cedex 07, France
Home page: `http://perso.ens-lyon.fr/fabio-lucio.toninelli`
`fltonine@ens-lyon.fr`

Summary. These notes are devoted to the statistical mechanics of directed polymers interacting with one-dimensional spatial defects. We are interested in particular in the situation where frozen disorder is present. These polymer models undergo a *localization/delocalization transition*. There is a large (bio)-physics literature on the subject since these systems describe, for instance, the statistics of thermally created loops in DNA double strands and the interaction between $(1 + 1)$-dimensional interfaces and disordered walls. In these cases the transition corresponds, respectively, to the DNA denaturation transition and to the wetting transition. More abstractly, one may see these models as random and inhomogeneous perturbations of renewal processes.

The last few years have witnessed a great progress in the mathematical understanding of the equilibrium properties of these systems. In particular, many rigorous results about the location of the critical point, about critical exponents and path properties of the polymer in the two thermodynamic phases (localized and delocalized) are now available.

Here, we will focus on some aspects of this topic—in particular, on the non-perturbative effects of disorder. The mathematical tools employed range from renewal theory to large deviations and, interestingly, show tight connections with techniques developed recently in the mathematical study of mean field spin glasses.

Contents

R. Kotecký (ed.), *Methods of Contemporary Mathematical Statistical Physics*, Lecture Notes in Mathematics 1970, DOI: 10.1007/978-3-540-92796-9_3,

1 Introduction and Motivations

Consider a Markov chain $\{S_n\}_{n\in\mathbb{N}}$ on some state space Ω, say, $\Omega = \mathbb{Z}^d$. We can unfold S along the discrete time axis, i.e., we can consider the sequence $\{(n, S_n)\}_{n\in\mathbb{N}}$ and interpret it as the configuration of a directed polymer in the space $\mathbb{N} \times \Omega$. In the examples which motivate our analysis, the discrete time is actually better interpreted as one of the space coordinates. The "directed" character of this polymer just refers to the fact that the first coordinate, n, is always increasing. In particular, the polymer can have no self-intersections. Some assumptions on the law of the Markov chain will be made in Section 2, where the model is defined precisely. Now let 0 be a specific point in Ω, and assume that the polymer receives a reward ϵ (or a penalty, if $\epsilon < 0$)

whenever $S_n = 0$, i.e., whenever it touches the *defect line* $\mathbb{N} \times \{0\}$. In other words, the probability of a configuration of $\{S_1, S_2, \ldots, S_N\}$ is modified by an exponential, Boltzmann-type factor

$$\exp\Big(\epsilon \sum_{n=1}^{N} \mathbf{1}_{\{S_n=0\}}\Big).$$

It is clear that if $\epsilon > 0$ contacts with the defect line are enhanced with respect to the $\epsilon = 0$ (or free) case, and that the opposite is true for $\epsilon < 0$. One can intuitively expect that the in the thermodynamic limit $N \to \infty$ a phase transition occurs: for $\epsilon > \epsilon_c$ the polymer stays close to the defect line essentially for every n, while for $\epsilon < \epsilon_c$ it is repelled by it and touches it only at a few places. This is indeed roughly speaking what happens, and the transition is given the name of *localization/delocalization transition*. We warn the reader that it is not true in general that the critical value is $\epsilon_c = 0$: if the Markov chain is transient, then $\epsilon_c > 0$, i.e., a strictly positive reward is needed to pin the polymer to the defect line (cf. Section 2.6).

A more interesting situation is that where the constant repulsion/attraction ϵ is replaced by a local, site-dependent repulsion/attraction ϵ_n. One can for instance consider the situation where ϵ_n varies periodically in n, but we will rather concentrate on the case where ϵ_n are independent and identically distributed (IID) random variables. We will see that, again, the transition exists when, say, the average ϵ of ϵ_n is varied. However, in this case the mechanism is much more subtle. This is reflected for instance in the counter-intuitive fact that ϵ_c may be negative: a globally repulsive defect line can attract the polymer! Presence of disorder opens the way to a large number of exciting questions, among which we will roughly speaking select the following one: how are the critical point and the critical exponents influenced by disorder?

There are several reasons to study disordered pinning models:

- there is a vast physics and bio-physics literature on the subject, with intriguing (but often contradictory) theoretical predictions and numerical/experimental observations. See also Section 2.6;
- they are interesting generalizations of classical renewal sequences. From this point of view they raise new questions and challenges, like the problem of the speed of convergence to equilibrium for the renewal probability in absence of translation invariance (cf. in particular Section 6);

- finally (and this is my main motivation) they are genuinely quenched-disordered systems where randomness has deep, non-perturbative effects. With respect to other systems like disordered ferromagnets or spin glasses, moreover, disordered pinning models have the advantage that their homogeneous counterparts are under full mathematical control. These models, therefore, turn out to be an ideal testing ground for theoretical physics arguments like the Harris criterion and renormalization group analysis.

It is also quite encouraging, from the point of view of mathematical physics, that rigorous methods have been able not only to confirm predictions made by theoretical physicists, but in some cases also to resolve controversies (it is the case for instance of the results in Section 5.6, which disprove some claims appeared previously in the physical literature).

1.1 A Side Remark on Literature and on the Scope of These Notes

A excellent recent introductory work on pinning models with quenched disorder (among other topics) is the book [22] by Giambattista Giacomin. In order to avoid the risk of producing a *résumé* of it, we have focussed on aspects which are not (or are only tangentially) touched in [22]. On the other hand, we will say very little about "polymer path properties", to which Chapters 7 and 8 of [22] are devoted. A certain degree of overlap is however inevitable, especially in the introductory sections 2 and 4; results taken from [22] will be often stated without proofs (unless they are essential in the logic of these notes).

We would also like to mention that some of the results of these notes apply also to a model much related to disordered pinning, namely *random heteropolymers (or copolymers) at selective interfaces*. It is the case, for instance, of the results of Sections 5.6 and 6. We have chosen to deal only with the pinning model for compactness of presentation, but we invite readers interested in the heteropolymer problem to look, for instance, at [11], [37], [22] and references therein.

2 The Model and its Free Energy

2.1 The Basic Renewal Process ("The Free Polymer")

Our starting point will be a renewal τ on the integers, $\tau := \{\tau_i\}_{i=0,1,2,\ldots}$, where $\tau_0 = 0$ and $\{\tau_i - \tau_{i-1}\}_{i\geq 1}$ are IID positive and integer-valued

random variables. The law of the renewal will be denoted by $\mathbf{P}$, and the corresponding expectation by $\mathbf{E}$. In terms of the "directed polymer picture" of the introduction, $\mathbf{P}$ is the law of the set τ of the points where the polymer touches the defect line, in absence of interaction: $\tau = \{n : S_n = 0\}$ (cf. also Section 2.6). We assume that $(\tau_i - \tau_{i-1})$ or, equivalently, τ_1 is $\mathbf{P}$-almost surely finite: if

$$K(n) := \mathbf{P}(\tau_1 = n), \tag{2.1}$$

this amounts to requiring $\sum_{n\in\mathbb{N}} K(n) = 1$. This, of course, implies that the renewal is recurrent: $\mathbf{P}$-almost surely, τ contains infinitely many points. A second assumption is that $K(.)$ has a power-like tail. More precisely, we require that

$$K(n) = \frac{L(n)}{n^{1+\alpha}} \text{ for every } n \in \mathbb{N}, \tag{2.2}$$

for some $\alpha \geq 0$ and a slowly varying function $L(.)$. We recall that a function $(0,\infty) \ni x \to L(x) \in (0,\infty)$ is said to be slowly varying at infinity if [8]

$$\lim_{x\to\infty} \frac{L(rx)}{L(x)} = 1 \tag{2.3}$$

for every $r > 0$. In particular, a slowly varying function diverges or vanishes at infinity slower than any power. The interested reader may look at [8] for properties and many interesting applications of slow variation. Of course, every positive function $L(.)$ having a non-zero limit at infinity is slowly varying. Less trivial examples are $L(x) = (\log(1+x))^\gamma$ for $\gamma \in \mathbb{R}$.

Observe that the normalization condition $\sum_{n\in\mathbb{N}} K(n) = 1$ implies that, if $\alpha = 0$, $L(.)$ must tend to zero at infinity (cf. also Section 2.6 below for an example).

It is important to remark that typical configurations of τ are very different according to whether α is larger or smaller than 1. Indeed the average distance between two successive points,

$$\mathbf{E}\,(\tau_i - \tau_{i-1}) = \sum_{n\in\mathbb{N}} nK(n), \tag{2.4}$$

is finite for $\alpha > 1$ and infinite for $\alpha < 1$. In standard terminology, τ is positively recurrent (i.e., τ occupies a finite fraction of $\mathbb{N}$) for $\alpha > 1$

and null-recurrent for $\alpha < 1$ (the density of τ in $\mathbb{N}$ is zero). This is a simple consequence of the classical renewal theorem [6, Chap. I, Th. 2.2], which states that

$$\lim_{n\to\infty} \mathbf{P}(n \in \tau) = \frac{1}{\sum_{n\in\mathbb{N}} nK(n)}. \tag{2.5}$$

The distinction $\alpha \gtrless 1$ plays an important role, especially in the behavior of the homogeneous pinning model (cf. Section 4). Later on we will see the emergence of an even more important threshold value: $\alpha_c = 1/2$.

Remark 2.1. For $\alpha = 1$, the question whether the renewal is positively or null recurrent is determined by the behavior at infinity of $L(.)$: from (2.5) we see that τ is finitely recurrent iff $\sum_n L(n)/n < \infty$. For instance, one has null recurrence if $L(.)$ has a positive limit at infinity.

2.2 The Model in Presence of Interaction

Now we want to introduce an interaction which favors the occurrence of a renewal at some points and inhibits it at others. To this purpose, let ω (referred to as *quenched randomness* or *random charges*) be a sequence $\{\omega_n\}_{n\in\mathbb{N}}$ of IID random variables with law $\mathbb{P}$. The basic assumption on ω_n, apart from the fact of being IID, is that $\mathbb{E}\,\omega_1 = 0$ and $\mathbb{E}\,\omega_1^2 = 1$. These are rather conventions than assumptions, since by varying the parameters β and h in Eq. (2.6) below one can effectively tune average and variance of the charges. To be specific, in these notes we will consider only two (important) examples: the Gaussian case $\omega_1 \stackrel{d}{=} \mathcal{N}(0,1)$ and the bounded case, $|\omega_1| \le C < \infty$. Many results are expected (or proven) to hold in wider generality and a few remarks in this direction are scattered throughout the notes.

We are now ready to define the free energy of our model: given $h \in \mathbb{R}$, $\beta \ge 0$ and $N \in \mathbb{N}$ let

$$F_N^\omega(\beta,h) := \frac{1}{N}\log Z_{N,\omega}(\beta,h) := \frac{1}{N}\log \mathbf{E}\left(e^{\sum_{n=1}^N (\beta\omega_n+h)\delta_n}\delta_N\right), \tag{2.6}$$

where for notational simplicity we put $\delta_n := \mathbf{1}_{\{n\in\tau\}}$, $\mathbf{1}_A$ being the indicator function of a set A. The *quenched average* of the free energy, or quenched free energy for short, is defined as

$$F_N(\beta,h) := \mathbb{E}F_N^\omega(\beta,h). \tag{2.7}$$

Note that the factor δ_N in (2.6) corresponds to imposing the boundary condition $N \in \tau$ (the boundary condition $0 \in \tau$ at the left border is

implicit in the law $\mathbf{P}$). One could equivalently work with free boundary conditions at N (i.e., replace δ_N by 1). The infinite-volume free energy would not change, but some technical steps in the proofs of some results would be (slightly) more involved.

We need also a notation for the Boltzmann-Gibbs average: given a realization ω of the randomness and a system size N, for a $\mathbf{P}$-measurable function $f(.)$ set

$$\mathbf{E}_{N,\omega}^{\beta,h}(f) := \frac{\mathbf{E}\left(f(\tau)\, e^{\sum_{n=1}^{N}(\beta\omega_n+h)\delta_n}\delta_N\right)}{Z_{N,\omega}(\beta,h)} \tag{2.8}$$

2.3 Existence and Non-negativity of the Free Energy

As usual in statistical mechanics, one is (mostly) interested in the thermodynamic limit (i.e., the limit $N \to \infty$). A classical question concerns the existence of the thermodynamic limit of the free energy, and its dependence on the realization of the randomness ω. In the context of the models we are considering, the answer is well established:

Theorem 2.2. *[22, Th. 4.1] If* $\mathbb{E}|\omega_1| < \infty$*, the limit*

$$F(\beta,h) := \lim_{N\to\infty} \frac{1}{N} \log Z_{N,\omega}(\beta,h) \tag{2.9}$$

exists for every $\beta \geq 0, h \in \mathbb{R}$ *and it is* $\mathbb{P}(\mathrm{d}\omega)$*-almost surely independent of* ω*.*

Of course, the limit does depend in general on the law $\mathbb{P}$ of the disorder.

Note that the only assumption on disorder, apart from the IID character of the charges, is finiteness of the first moment, so that existence and self-averaging of the infinite-volume free energy holds in much wider generality than in the cases of Gaussian or bounded disorder we are considering here.

Some properties of the free energy come essentially for free: in particular, $F(\beta,h)$ is convex in (β,h), non-decreasing in h, continuous everywhere and differentiable almost everywhere as a consequence of convexity. Another easy fact is that the sequence $\{N\, F_N(\beta,h)\}_{N\in\mathbb{N}}$ is

super-additive: for every $N, M \in N$, one has $(N+M)F_{N+M}(\beta,h) \geq NF_N(\beta,h) + MF_M(\beta,h)$. This is easily proven:

$$\begin{aligned}(N+M)F_{N+M}(\beta,h) &= \mathbb{E}\log\mathbf{E}\left(e^{\sum_{n=1}^{N+M}(\beta\omega_n+h)\delta_n}\delta_{N+M}\right)\\ &\geq \mathbb{E}\log\mathbf{E}\left(e^{\sum_{n=1}^{N}(\beta\omega_n+h)\delta_n}\delta_N e^{\sum_{n=N+1}^{N+M}(\beta\omega_n+h)\delta_n}\delta_{N+M}\right)\\ &= NF_N(\beta,h) + MF_M(\beta,h), \quad (2.10)\end{aligned}$$

where in the last step we used invariance of $\mathbb{P}$ with respect to left shifts and the renewal property of $\mathbf{P}$. It is a standard fact that super-additivity implies

$$F(\beta,h) \geq F_N(\beta,h) \text{ for every } N \in \mathbb{N}. \quad (2.11)$$

2.4 Contact Fraction and Critical Point

As we already mentioned, the interest in this class of models is mainly due to the fact that they show a so-called *localization-delocalization* transition. This is best understood in view of the elementary bound $F(\beta,h) \geq 0$. This positivity property is immediate to prove:

$$F_N(\beta,h) \geq \frac{1}{N}\mathbb{E}\log\mathbf{E}\left(e^{\sum_{n=1}^{N}(\beta\omega_n+h)\delta_n}\mathbf{1}_{\{\tau_1=N\}}\right) = \frac{h}{N} + \frac{1}{N}\log K(N) \quad (2.12)$$

and the claimed non-negativity in the limit follows from (2.2). Recalling that $F(\beta,h)$ is non-decreasing in h, for a given β the localization/delocalization critical point is defined to be

$$h_c(\beta) := \sup\{h : F(\beta,h) = 0\} \quad (2.13)$$

and the function $\beta \to h_c(\beta)$ is referred to as the *critical line*. The region of parameters

$$\mathcal{L} := \{(\beta,h) : \beta \geq 0, h > h_c(\beta)\}$$

and

$$\mathcal{D} := \{(\beta,h) : \beta \geq 0, h \leq h_c(\beta)\}$$

are referred to as *localized* and *delocalized* phases, respectively. Since level sets of a convex function are convex, $\mathcal{L}$ is a convex set and the function $h_c(.) : [0,\infty) \ni \beta \to h_c(\beta)$ is concave. The reason for the names "localized" and "delocalized" can be understood looking at the so-called *contact fraction* ℓ_N, defined through

$$\ell_N := \frac{|\tau \cap \{1,\ldots,N\}|}{N} \quad (2.14)$$

and taking values between 0 and 1 (as usual, $|A|$ denotes the cardinality of a set A). It is immediate to check that

$$\partial_h F_N^\omega(\beta, h) = \mathbf{E}_{N,\omega}^{\beta,h}(\ell_N) \tag{2.15}$$

and, by standard arguments based on convexity, this equality survives in the thermodynamic limit whenever the free energy is differentiable:

$$\lim_{N\to\infty} \mathbf{E}_{N,\omega}^{\beta,h}(\ell_N) \overset{a.s.}{=} \partial_h F(\beta, h) \quad \text{for every } h \text{ such that}$$
$$\partial_h^+ F(\beta, h) = \partial_h^- F(\beta, h). \tag{2.16}$$

We have already mentioned that differentiability holds for Lebesgue-almost every value of h. However, much more than this is true: as it was proven in [25], differentiability (actually, infinite differentiability) in h holds whenever $h > h_c(\beta)$. We can therefore conclude the following: for $h < h_c(\beta)$ (or for $h \le h_c(\beta)$ if $F(\beta, h)$ is differentiable at $h_c(\beta)$) the thermal average of the contact fraction tends for to zero for $N \to \infty$ (almost surely in the disorder), while for $h > h_c(\beta)$ it tends to $\partial_h F(\beta, h) > 0$. The average contact fraction plays the role of an order parameter, like the spontaneous magnetization in the Ising model, which is zero above the critical temperature and positive below it.

Actually, much more refined statements about the behavior of the contact fraction in the two phases are available. In particular:

- for statements concerning the localized phase we refer to [25]. There, it is proven that, roughly speaking, not only typical configurations τ have a number
$$N\,\ell_N \sim N\,\partial_h F(\beta, h)$$
of points, but also that these points are rather uniformly distributed in $\{1, \dots, N\}$: long gaps between them are exponentially suppressed, and the largest gap is of order $\log N$ (cf. Theorem 6.3 below);
- for $h < h_c(\beta)$ we refer to [24] and [22, Ch. 8], where it is proven that ℓ_N is typically at most of order $(\log N)/N$.

In this sense, if one goes back to the pictorial image of τ as the set of points of polymer-defect contact, one sees that the definition of (de)-localization in terms of free energy, as given above, does indeed correspond to the intuitive idea in terms of path properties: in $\mathcal{L}$ the polymer stays at distance $O(1)$ from the defect, while in $\mathcal{D}$ it wanders away from it and touches it only a small (at most $\log N$) number of times.

The reader should remark that we have made no conclusive statement about the behavior of the contact fraction at $h_c(\beta)$, since we have

not attacked yet the very important question of the regularity of the free energy at the critical point. This will be the subject of Sections 4 and 5.

2.5 Quenched versus Annealed Free Energy

Inequality (2.12) is a good example of how selecting a particular subset of configurations (in that case, those for which $\tau_1 = N$) provides useful free energy lower bounds. For more refined results in this direction we refer to [5] and [22, Sec. 5.2]. There, this technique is employed to prove that $h_c(\beta)$ is strictly decreasing as a function of β which implies in particular that, since $h_c(.)$ is concave, $h_c(\beta)$ tends to $-\infty$ for $\beta \to \infty$. This corresponds to the a priori non-intuitive fact that, as mentioned in the introduction, even if the charges are on average repulsive the defect line can pin the polymer. This is purely an effect of spatial inhomogeneities due to disorder: for β large, it is convenient for the polymer to touch the defect line in correspondence of attractive charges, where it gets a reward $\beta\omega_n + h >> 1$, while the entropic cost of avoiding the repulsive charges is independent of β. Free energy lower bounds were obtained also in the study of a different model, the *heteropolymer at a selective interface*, in [10].

Free energy upper bounds are on the other hand more subtle to get. An immediate one can be however obtained by a simple application of Jensen's inequality:

$$F_N(\beta,h) \le \frac{1}{N}\log \mathbb{E} Z_{N,\omega}(\beta,h) = \frac{1}{N}\log \mathbf{E}\left(e^{\sum_{n=1}^N (h+\log M(\beta))\delta_n}\delta_N\right)$$
$$= F_N(0, h+\log M(\beta)) =: F_N^a(\beta,h), \quad (2.17)$$

where $M(\beta) := \mathbb{E}\, e^{\beta\omega_1}$. In particular, $\log M(\beta) = \beta^2/2$ in the case of Gaussian disorder. $F^a(\beta,h) := F(0, \beta + \log M(\beta))$ is referred to as *annealed free energy*, and we see that it is just the free energy of the homogeneous system (with the same choice of $K(.)$) computed for a shifted value of h. The physical interpretation of the annealed free energy is clear: since configurations of ω and τ are averaged on the same footing, it corresponds to a system where impurities can thermalize on the same time-scales as the "polymer degrees of freedom" (i.e., τ). This is not the physical situation one wishes to study (quenched disorder corresponds rather to impurities which are frozen, or which can evolve only on time-scales which are so long that they can be considered as infinite from the experimental point of view). All the same, the

information provided by (2.17) is not at all empty. Define first of all the *annealed critical point* as

$$h_c^a(\beta) := \sup\{h : F^a(\beta, h) = 0\} = h_c(0) - \log M(\beta). \tag{2.18}$$

Thanks to (2.17) and (2.13), one has immediately

$$h_c(\beta) \geq h_c(0) - \log M(\beta), \tag{2.19}$$

a bound which, as will be discussed in Section 5.3, is optimal for $\alpha < 1/2$ and β small.

2.6 Back to Examples and Motivations

Typical examples of renewal sequences satisfying (2.1), (2.2) are the following. Let $\{S_n\}_{n\geq 0}$ be the simple random walk (SRW) on $\mathbb{Z}$, with law $\mathbf{P}^{SRW}$ and $S_0 := 0$, i.e., $\{S_n - S_{n-1}\}_{n\in\mathbb{N}}$ are IID symmetric random variables with values in $\{-1, +1\}$. Then, it is known that [19] $\tau := \{n \in \mathbb{N} : S_{2n} = 0\}$ is a null-recurrent renewal sequence such that the law of τ_1 satisfies (2.2) with $\alpha = 1/2$ and $L(.)$ asymptotically constant. The reason why one looks only at even values of n in the definition of τ in this case is due just to the periodicity of the SRW. If instead one takes the SRW on $\mathbb{Z}^2$, then τ (defined exactly as above) is always a null-recurrent renewal but in this case $\alpha = 0$ and $L(n) \sim c/(\log n)^2$ [34]. Note that in this case, the presence of the slowly varying function $L(.)$ is essential in making $K(.)$ summable.

What happens in the case of the SRW on $\mathbb{Z}^d$ when $d \geq 3$? This example does not fall directly into the class we are considering since this process is transient, and therefore the set τ of its returns to zero is a transient renewal sequence. However this is not too bad. Indeed, suppose more generally that one is given $K(.)$ which satisfies (2.2) but such that $\Sigma := \sum_{n\in\mathbb{N}} K(n) < 1$, i.e., $K(.)$ is a sub-probability on $\mathbb{N}$. Then, one may define $\hat{K}(n) := K(n)/\Sigma$ which is obviously a probability. It is easy to realize from Eq. (2.8) that the Gibbs measure (and free energy) of the model defined starting from $K(.)$ is the same as that obtained starting from $\hat{K}(.)$, provided that h is replaced by $h + \log \Sigma$. The case where τ are the zeros of the SRW on $\mathbb{Z}^d$ with $d \geq 3$ can then be included in our discussion: Eq. (2.2) holds with $\alpha = d/2 - 1$ and $L(.)$ asymptotically constant. In the following we will therefore always assume, without loss of generality, that τ is recurrent.

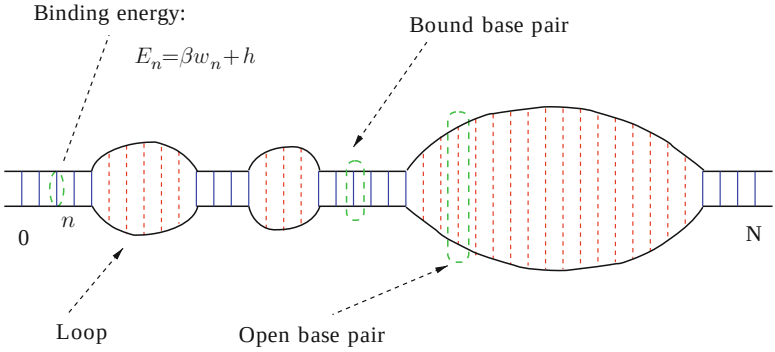
We conclude this section by listing a couple of examples of (bio)-physical situations where disordered pinning models are relevant:

- *Wetting of* $(1+1)$*-dimensional disordered substrates* [17] [21]. Consider a two-dimensional system at a first order phase transition, e.g., the $2d$-Ising model at zero magnetic field and $T < T_c$, or a liquid-gas system on the coexistence line. Assume that the system is enclosed in a square box with boundary conditions imposing one of the two phases along the bottom side of the box and the other phase along the other three sides. For instance, for the Ising model one can impose $+$ boundary conditions (b.c.) along the bottom side and $-$ b.c. along the other ones; for the liquid-gas model, one imposes that the bottom of the box is in contact with liquid and that side and top walls are in contact with gas. Then, there is necessarily an interface joining the two bottom corners of the box and separating the two phases. At very low temperature, it is customary to describe this interface as a one-dimensional symmetric random walk (not necessarily the SRW) conditioned to be non-negative, the non-negativity constraint reflecting the fact that the interface cannot exit the box. The directed character of the random walk implies in particular that one is neglecting the occurrence of bubbles or overhangs in the interface. An interesting situation occurs when the bottom wall is "dirty" and at each point has a random interaction with the interface: at some points the wall prefers to be in contact with the gas (or $-$ phase), and therefore tries to pin the interface, while at other points it prefers contact with the liquid (or $+$ phase) and repels the interface. Of course, this non-homogeneous interaction is encoded in the charges ω_n. In this context, the (de)-localization transition is called *wetting transition*. This denomination is clear if we think of the liquid-gas model: the localized phase corresponds to an interface which remains at finite distance from the wall (the wall is dry), while in the delocalized phase there are few interface-wall contacts and the height of the liquid layer on the wall diverges in the thermodynamic limit: the wall is wet. It is known that, in great generality [19], the law of the first return to zero of a one-dimensional random walk conditioned to be non-negative is of the form (2.2) with $\alpha = 1/2$ and $L(.)$ asymptotically constant (this process is transient but this fact is not so relevant, in view of the discussion at the beginning of the present section).
- *Formation of loops under thermal excitation and denaturation of DNA molecules* in the Poland-Scheraga (PS) approximation [15]. Neglecting its helical structure, the DNA molecule is essentially a

double strand of complementary units, called "bases". Upon heating, the bonds which keep base pairs together can break and the two strands can partly or entirely separate (cf. figure below). This separation, or denaturation, can be described in the context of our disordered pinning models. The set τ represents the set of bases whose bond is not broken. In the localized phase τ contains $O(N)$ points (N being interpreted here as the total DNA length), i.e., corresponds to the phase where the two strands are still essentially tightly bound. In the delocalized (or denaturated) phase, on the contrary, only few bases pairs are bound. In formulating the PS model, one usually takes a value $\alpha \simeq 2.12$ (cf. [35] for a justification of this choice) and (in our notations, which are not necessarily those of the literature on the PS model)

$$L(n) = \sigma \quad \text{for} \quad n \geq 2,$$

where σ (the *cooperativity parameter*) is a small number, usually of the order 10^{-5}, while $L(1)$ is fixed by the normalization condition $\sum_{n\in\mathbb{N}} K(n) = 1$. Quenched disorder corresponds here to the fact that bases of the different types are placed inhomogeneously along the DNA chain. We refer to [22, Section 1.4] for a very clear introduction to the denaturation problem and the Poland-Scheraga model. Here we wish to emphasize only that the renewal process τ described by such a $K(.)$ is not in general the set of returns of a Markov chain, as it happens for instance in the case of the wetting model described above.



3 The Questions We are Interested in

The main questions which will be considered in these notes are the following:

1. When is the annealed bound (2.17) a good one, i.e., when are quenched and annealed systems similar? We will see that quenched and annealed free energies never coincide, except in the (trivial) case where the annealed free energy is zero (i.e., the annealed model is delocalized). However, this does not mean that the solution of the annealed system gives no information about the quenched one. For instance we will show that, for $\alpha < 1/2$ and weak enough disorder, the quenched critical point coincides with the annealed one. This will be discussed in Section 5.3.
2. What is the order of the transition? Critical exponents (in particular, the *specific heat exponent*, cf. next section) can be exactly computed for the homogeneous model. The Harris criterion predicts that for small β critical exponents are those of the $\beta = 0$ (or annealed) model if $\alpha < 1/2$, and are different if $\alpha > 1/2$. This is the question of *disorder relevance*, discussed in Sections 5.3–5.6.
3. Truncated correlations functions are known to decay exponentially at large distance, in the localized phase. What is the behavior of the correlation length when the transition is approached? We will see that, due to the presence of quenched disorder, one can actually define two different correlation lengths. In specific cases, we will identify these correlation lengths and give bounds on the critical exponents which govern their divergence at $h_c(\beta)$.

4 The Homogeneous Model

In absence of disorder ($\beta = 0$) the model is under full mathematical control; in particular, critical point and the order of the transition can be computed exactly. In this section, we collect a number of known results, referring to [22, Chapter 2] for their proofs.

The basic point is that the free energy $F(0, h)$ is determined as follows [26, Appendix A]: if the equation

$$\sum_{n\in\mathbb{N}} e^{-bn} K(n) = e^{-h} \tag{4.1}$$

has a positive solution $b = b(h) > 0$ then $F(0, h) = b(h)$. Otherwise, $F(0, h) = 0$. From this (recall the normalization condition $\sum_{n\in\mathbb{N}}$

$K(n) = 1$), one finds immediately that $h_c(0) = 0$. The behavior of the free energy in the neighborhood of $h_c(0)$ can be also obtained from (4.1). Care has to be taken since a naive expansion of left- and right-hand sides of (4.1) for b and h small does not work in general. However, this analysis can be performed without much difficulty and one can prove the following:

Theorem 4.1. *[22, Th. 2.1]*

1. *If $\alpha = 0$, $F(0,h)$ vanishes faster than any power of h for $h \searrow 0$.*
2. *If $0 < \alpha < 1$ then for $h > 0$*

$$F(0,h) = h^{1/\alpha} \hat{L}(1/h), \tag{4.2}$$

where $\hat{L}(.)$ is the slowly varying function

$$\hat{L}(1/h) = \left(\frac{\alpha}{\Gamma(1-\alpha)} \right)^{1/\alpha} h^{-1/\alpha} R_\alpha(h) \tag{4.3}$$

and $R_\alpha(.)$ is asymptotically equivalent to the inverse of the map $x \to x^\alpha L(1/x)$.

3. *If $\alpha = 1$ and $\sum_{n\in\mathbb{N}} nK(n) = \infty$ then $F(0,h) = h\,\hat{L}(1/h)$ for some slowly varying function $\hat{L}(.)$ which vanishes at infinity.*
4. *If $\sum_{n\in\mathbb{N}} nK(n) < \infty$ (in particular, if $\alpha > 1$)*

$$F(0,h) \overset{h\searrow 0}{\sim} \frac{h}{\sum_{n\in\mathbb{N}} nK(n)}. \tag{4.4}$$

In particular, note that in the situation (4), i.e., if τ is positively recurrent under $\mathbf{P}$, the transition is of first order: the free energy is not differentiable at $h_c(0) = 0$, i.e., the average contact fraction has a finite jump in the thermodynamic limit. This is analogous to what happens for the Ising model in dimension $d \geq 2$: if $T < T_c$ and one varies the magnetic field H from 0^- to 0^+, the spontaneous magnetization has a positive jump and the free energy is not differentiable. The transition is, on the other hand, continuous (*at least of second order*) if $\mathbf{P}$ is the law of a null-recurrent renewal τ and it becomes smoother as α decreases. In thermodynamical language, one can say that the delocalization transition is of k^{th} order ($F(\beta,.)$ is of class C^{k-1} but not of class C^k) for $\alpha \in (1/k, 1/k-1)$ and of infinite order for $\alpha = 0$.[1]

[1] In order to decide between k^{th} and $(k+1)^{th}$ order for $\alpha = 1/k$ one needs to look also at the slowly varying function $L(.)$, as is already clear from points (3) and (4) in the case of $k = 1$. In any case, the precise statement is that of Theorem 4.1.

In the physics literature one introduces usually the *specific heat critical exponent* ν as[2]

$$\nu = 2 - \lim_{h \searrow h_c(\beta)} \frac{\log F(\beta, h)}{\log(h - h_c(\beta))} \tag{4.5}$$

(provided the limit exists) and of course ν can depend on β. From Theorem 4.1 we see that, in absence of disorder,

$$\nu(\beta = 0) = 2 - \max(1, 1/\alpha). \tag{4.6}$$

In particular, note that $\nu(\beta = 0) > 0$ as soon as $\alpha > 1/2$ (this observation will become interesting in the light of the results of Section 5.6).

5 Relevance or Irrelevance of Disorder?

We have just seen that the phase transition of the homogeneous pinning model can be of any given order - from first to infinite - depending on the choice of $K(.)$ in (2.2) and, in particular, on the value of α. In this section we discuss the effect of disorder on the transition and we are primarily interested in the question of disorder relevance. There are actually two distinct (but inter-related) aspects in this question:

Q1 does an arbitrarily small quantity of disorder change the critical exponent ν (i.e., the order of the transition)?
Q2 does the quenched critical point differ from the annealed one for very weak disorder?

One expects the answer to both questions to be "no" if $\alpha < 1/2$ and "yes" if $\alpha > 1/2$, while the case $\alpha = \alpha_c = 1/2$ is more subtle and not clear even heuristically [17, 21] (see, however, Theorem 5.5).

The plan is the following: we will first of all (Section 5.1) make a non-rigorous computation, in the spirit of the Harris approach [33], which shows why the watershed value for α, distinguishing between relevance and irrelevance, is expected to be $\alpha_c = 1/2$, i.e., the value for which the critical exponent ν vanishes for the homogeneous model (cf. (4.6)). Next, in Section 5.2 we prove an upper bound for the free energy which strictly improves the annealed bound (2.17). In the proof of this bound we introduce the technique of *interpolation*, by now classical in spin glass theory but sort of new in this context. We would like to emphasize

[2] The symbol ν for the specific heat exponent is not standard in the literature, but we have already used the letter α for another purpose. The same remark applies to the symbols we use for other critical exponents.

that interpolation (and replica coupling, cf. Section 5.5) techniques have proven recently to be extremely powerful in the analysis of mean field spin glass models, cf. for instance [32], [1], [39], while their relevance in the domain of disordered pinning model had not been realized clearly so far.

As a byproduct, our new upper bound partially justifies the heuristic expansion of Section 5.1. The question of relevance is taken up more seriously in Sections 5.3 to 5.6. In the former we will see, among other results, that answers to both **Q1** and **Q2** are actually "no" for $\alpha < \alpha_c$. In the latter, on the other hand, we show that critical exponents *are* modified by disorder for $\alpha > \alpha_c$: in particular, we will see that $\nu \leq 0$ whenever $\beta > 0$.

In the whole of Section 5 we consider only the case of Gaussian disorder. This allows for technically simpler proofs, but results can be generalized for instance to the bounded disorder case.

5.1 Harris Criterion and the Emergence of $\alpha_c = 1/2$

Let us note for clarity that, putting together the discussion of Section 4 and Eq. (2.18), in the Gaussian case the annealed critical point equals $h_c^a(\beta) = -\beta^2/2$. The first step of our heuristic argument is rigorous and, actually, an immediate identity:

$$F_N(\beta, h) = F_N^a(\beta, h) + \frac{1}{N}\mathbb{E}\log\left\langle e^{\sum_{n=1}^N (\beta\omega_n - \beta^2/2)\delta_n}\right\rangle_{N, h-h_c^a(\beta)}, \quad (5.1)$$

where $\langle . \rangle_{N,h} := \mathbf{E}_{N,0}^{0,h}(.)$ is just the Boltzmann average for the homogeneous system (cf. Eq. (2.8)). Identity (5.1) can be rewritten in a more suggestive way if we recall the last equality in (2.17) and we let $h = h_c^a(\beta) + \Delta$ with $\Delta \geq 0$:

$$\begin{aligned} F_N(\beta, h_c^a(\beta) + \Delta) &= F_N(0, \Delta) + R_{N,\Delta}(\beta) := \\ &= F_N(0, \Delta) + \frac{1}{N}\mathbb{E}\log\left\langle e^{\sum_{n=1}^N (\beta\omega_n - \beta^2/2)\delta_n}\right\rangle_{N,\Delta}. \end{aligned} \quad (5.2)$$

Irrelevance of disorder amounts to the fact that, for β sufficiently small, the "error term" $R_{N,\Delta}(\beta)$ is negligible with respect to the "main term" $F_N(0, \Delta)$. As we will see, the question is subtle since we are interested in both Δ and β small, and the two limits do not in general commute. For the moment, let us proceed without worrying about rigor and let us expand naively $R_{N,\Delta}(\beta)$ for β small and Δ, N fixed:

$$\left\langle e^{\sum_{n=1}^{N}(\beta\omega_n-\beta^2/2)\delta_n}\right\rangle_{N,\Delta} = 1+\sum_{n=1}^{N}(\beta\omega_n-\beta^2/2)\left\langle\delta_n\right\rangle_{N,\Delta}+ \\ +\frac{\beta^2}{2}\sum_{n,m=1}^{N}\omega_n\omega_m\left\langle\delta_n\delta_m\right\rangle_{N,\Delta}+O(\beta^3). \quad (5.3)$$

Expanding the logarithm and using the fact that $\mathbb{E}\,\omega_n = 0$ and $\mathbb{E}(\omega_n\omega_m) = \mathbf{1}_{\{n=m\}}$ one has, always formally,

$$R_{N,\Delta}(\beta) = -\frac{\beta^2}{2N}\sum_{n=1}^{N}\left(\left\langle\delta_n\right\rangle_{N,\Delta}\right)^2+O(\beta^3). \quad (5.4)$$

In the limit $N\to\infty$ one has by definition of the homogeneous model

$$\lim_{N\to\infty}\left\langle\ell_N\right\rangle_{N,\Delta} = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\left\langle\delta_n\right\rangle_{N,\Delta} = \partial_\Delta F(0,\Delta).$$

Since $\left\langle\delta_n\right\rangle_{N,\Delta}$ should not depend on n as soon as $1\ll n\ll N$, one can expect (actually, this can be proven without much difficulty) that

$$\lim_{\substack{N\to\infty\\ n/N\to m\in(0,1)}}\left\langle\delta_n\right\rangle_{N,\Delta} = \partial_\Delta F(0,\Delta). \quad (5.5)$$

In conclusion, we find

$$F(\beta,h_c^a(\beta)+\Delta) = F(0,\Delta)-\frac{\beta^2}{2}(\partial_\Delta F(0,\Delta))^2+O(\beta^3). \quad (5.6)$$

Even without trying (for the moment) to justify this expansion or to look more closely at the Δ-dependence of the error term $O(\beta^3)$, we can extract something important from Eq. (5.6). We know from Theorem 4.1 that, for $\alpha<1$ and $\Delta>0$ small, $F(0,\Delta)\simeq\Delta^{1/\alpha}$ which implies (cf. the proof of Eq. (5.19) for details) that $\partial_\Delta F(0,\Delta)\simeq\Delta^{1/\alpha-1}$. Then we see immediately that, indeed, for $\alpha<1/2$

$$\frac{\beta^2}{2}(\partial_\Delta F(0,\Delta))^2\ll F(0,\Delta) \quad (5.7)$$

if Δ and β are small. In terms of the Harris criterion, disordered is said to be irrelevant in this case and one can hope that the expansion can be actually carried on at higher orders. For $1/2<\alpha<1$, however, this is false: even if β is small, choosing Δ sufficiently close to zero the left-hand side of (5.7) is much larger than the right-hand side. This means

that "disorder is relevant" and the small-disorder expansion breaks down immediately. The same holds for $\alpha \geq 1$, when $F(0, \Delta) \simeq \Delta$ and $\partial_\Delta F(0, \Delta) \sim const.$ The threshold value $\alpha_c = 1/2$ is clearly a "marginal case" where relevance or irrelevance of disorder cannot be decided (even on heuristic grounds) by a naive expansion in β.

The rest of this section will be devoted to give rigorous bases to this suggestive picture. As a byproduct we will learn something interesting for the case $1/2 < \alpha < 1$: while disorder is relevant and changes the exponent ν, it modifies the transition only "very close" to the critical point (cf. Theorem 5.3).

5.2 A Rigorous Approach: Interpolation and an Improvement upon Annealing

In Section 2.5 we saw that a simple application of Jensen's inequality implies $F(\beta, h) \leq F^a(\beta, h)$. Here we wish to show that this inequality is strict as soon as disorder is present ($\beta > 0$) and the annealed system is localized. Moreover, we will partly justify the small-β expansion of Section 5.1 for $\alpha < 1/2$, showing that it provides an upper bound for the quenched free energy.

More precisely:

Theorem 5.1. *[43, Th. 2.6] For every $\beta > 0$, $\alpha \geq 0$ and $\Delta > 0$*

$$F(\beta, h_c^a(\beta) + \Delta) \leq \inf_{0 \leq q \leq \Delta/\beta^2} \Big(\frac{\beta^2 q^2}{2} + F(0, \Delta - \beta^2 q) \Big) < \\ < F(0, \Delta) = F^a(\beta, h). \quad (5.8)$$

In particular, if $0 \leq \alpha < 1/2$ there exist constants $\beta_0 > 0, \Delta_0 > 0$ such that

$$F(\beta, h_c^a(\beta) + \Delta) \leq F(0, \Delta) - \frac{\beta^2}{2} \left(\partial_\Delta F(0, \Delta)\right)^2 (1 + O(\beta^2)) \quad (5.9)$$

for $\beta \leq \beta_0, \Delta \leq \Delta_0$, where $O(\beta^2)$ is does not depend on Δ. On the other hand, if $\beta = 0$ or $\Delta \leq 0$, then $F(\beta, h_c^a(\beta) + \Delta) = F^a(\beta, h_c^a(\beta) + \Delta)$.

About the possibility of pushing the upper bound (5.9) to order higher than β^2 see Remark 3.1 in [43]. It is obvious that (5.9) cannot hold for $\alpha > 1/2$ since, as already observed after Eq. (5.7), the right-hand side is negative for Δ sufficiently small.

Readers familiar with mean field spin glass models will remark a certain similarity between the variational bound (5.8) and the

"replica symmetric" variational bound [30] for the free energy of the Sherrington-Kirkpatrick model. However, we do not see a natural way to generalize (5.8) to include "replica symmetry breaking" in analogy with [29] [1].

Proof of Theorem 5.1. The proof is rather instructive because it allows us to introduce the technique of "interpolation", which will play a major role in the next subsection. We start from identity (5.2) and, for $\Delta > 0, q \in \mathbb{R}$ and $0 \leq t \leq 1$, we define

$$R_{N,\Delta}(t,\beta,q) := \frac{1}{N}\mathbb{E}\log\left\langle e^{\sum_{n=1}^{N}[\beta\sqrt{t}\omega_n - t\beta^2/2 + \beta^2 q(t-1)]\delta_n}\right\rangle_{\Delta,N}. \quad (5.10)$$

In spin glass language, this would be called an "interpolating free energy", since by varying the parameter t it relates in a smooth way the quantity we wish to estimate at $t = 1$,

$$R_{N,\Delta}(t=1,\beta,q) = R_{N,\Delta}(\beta) \quad (5.11)$$

to something easy at $t = 0$:

$$R_{N,\Delta}(t=0,\beta,q) = F_N(0,\Delta - \beta^2 q) - F_N(0,\Delta). \quad (5.12)$$

A priori, there is no reason why $R_{N,\Delta}(t,\beta,q)$ should be any easier to compute for $0 < t < 1$ than for $t = 1$. What helps us is that the t-derivative of $R_{N,\Delta}(t,\beta,q)$ can be bounded above by throwing away a (complicated) term which, luckily, has a negative sign. To see this we need first of all manageable notations and we will set

$$\langle g(\tau)\rangle_{N,\Delta,t} := \frac{\left\langle g(\tau)e^{\sum_{n=1}^{N}[\beta\sqrt{t}\omega_n - t\beta^2/2 + \beta^2 q(t-1)]\delta_n}\right\rangle_{\Delta,N}}{\left\langle e^{\sum_{n=1}^{N}[\beta\sqrt{t}\omega_n - t\beta^2/2 + \beta^2 q(t-1)]\delta_n}\right\rangle_{N,\Delta}} \quad (5.13)$$

for every measurable function $g(\tau)$. We find then

$$\frac{\mathrm{d}}{\mathrm{d}t}R_{N,\Delta}(t,\beta,q) = \frac{\beta^2}{N}\left(-\frac{1}{2}+q\right)\sum_{m=1}^{N}\mathbb{E}\,\langle\delta_m\rangle_{N,\Delta,t} + \\ + \frac{\beta}{2\sqrt{tN}}\sum_{m=1}^{N}\mathbb{E}\,\omega_m\,\langle\delta_m\rangle_{N,\Delta,t}. \quad (5.14)$$

The last term of (5.14) can be rewritten using the Gaussian integration by parts formula

$$\mathbb{E}\,(\omega f(\omega)) = \mathbb{E}f'(\omega), \quad (5.15)$$

which holds (if ω is a Gaussian random variable $\mathcal{N}(0,1)$) for every differentiable function $f(.)$ such that $\lim_{|x|\to\infty} \exp(-x^2/2)f(x) = 0$. In our case, the function f is of course $\langle \delta_m \rangle_{N,\Delta,t}$ and one finds

$$\frac{\beta}{2\sqrt{t}N} \sum_{m=1}^{N} \mathbb{E}\, \omega_m \langle \delta_m \rangle_{N,\Delta,t} = \frac{\beta^2}{2N} \sum_{m=1}^{N} \mathbb{E} \left(\langle \delta_m \rangle_{N,\Delta,t} - \left(\langle \delta_m \rangle_{N,\Delta,t} \right)^2 \right). \tag{5.16}$$

The positive term comes from the differentiation of the numerator of $\langle \delta_m \rangle_{N,\Delta,t}$ (recall the definition (5.13)) and the negative one from the denominator, and we used the obvious $\delta_m = (\delta_m)^2$. Putting together Eqs. (5.14) and (5.16) one has therefore

$$\frac{\mathrm{d}}{\mathrm{d}t} R_{N,\Delta}(t,\beta,q) = \frac{\beta^2 q^2}{2} - \frac{\beta^2}{2N} \sum_{n=1}^{N} \mathbb{E} \left\{ \left(\langle \delta_n \rangle_{N,\Delta,t} - q \right)^2 \right\} \leq \frac{\beta^2 q^2}{2}. \tag{5.17}$$

At this point we are done: we integrate on t between 0 and 1 inequality (5.17), we recall the boundary conditions (5.12) and (5.11) and we get

$$R_{N,\Delta}(\beta) \leq F_N(0, \Delta - \beta^2 q) - F_N(0,\Delta) + \frac{\beta^2 q^2}{2}. \tag{5.18}$$

Together with Eq. (5.1), taking $N \to \infty$ limit and minimizing over q proves (5.8). Let us remark that minimizing over $q \in \mathbb{R}$ or on $0 \leq q \leq \Delta/\beta^2$ is clearly equivalent. The strict inequality in (5.8) is just due to the fact that the derivative with respect to q of the quantity to be minimized, computed at $q = 0$, is negative.

The expansion (5.9) is just a consequence of (5.8). Remark first of all that, at the lowest order in β, the minimizer in (5.8) is $q = q_\Delta := \partial_\Delta F(0,\Delta)$. Then, from identity (4.1) one finds that there exist slowly varying functions $L^{(i)}(.), i = 1,2$ such that for $\alpha < 1/2$ and $\Delta > 0$

$$\partial_\Delta F(0,\Delta) = \Delta^{(1-\alpha)/\alpha} L^{(1)}(1/\Delta),\ \partial_\Delta^2 F(0,\Delta) = \Delta^{(1-2\alpha)/\alpha} L^{(2)}(1/\Delta). \tag{5.19}$$

Let us show for instance the first equality. Differentiating both sides of (4.1) with respect to Δ one finds

$$\partial_\Delta F(0,\Delta) = \frac{e^{-\Delta}}{\sum_{n\in\mathbb{N}} n^{-\alpha} L(n) \exp(-F(0,\Delta)n)}. \tag{5.20}$$

Using Theorems A.1 and A.2 one has then, for $\Delta \to 0$ (i.e., for $F(0,\Delta) \to 0$)

$$\partial_\Delta F(0,\Delta) \overset{\Delta\searrow 0}{\sim} \frac{\Gamma(2-\alpha)L(1/F(0,\Delta))}{(1-\alpha)F(0,\Delta)^{1-\alpha}} \tag{5.21}$$

which, together with (4.2), proves the first equality in (5.19) for a suitable $L^{(1)}(.)$. Note, by the way, that thanks to (5.19) one has $q_\Delta < \Delta/\beta^2$ for Δ, β sufficiently small (and $\alpha < 1/2$, of course). Another consequence of (5.19) is that $\partial_\Delta^2 F(0,\Delta)$ is bounded above by a finite constant C for, say, $\Delta \leq 1$. Then, a Taylor expansion gives

$$F(0,\Delta-\beta^2 q_\Delta) \leq F(0,\Delta) - \beta^2(\partial_\Delta F(0,\Delta))^2 + C\beta^4(\partial_\Delta F(0,\Delta))^2,$$

whence Eq. (5.9).

Finally, the statement for $\beta = 0$ or $\Delta \leq 0$ is trivial: for $\beta = 0$ there is no disorder to distinguish between quenched an annealed free energies, and for $\Delta \leq 0$ one has $F^a(\beta, h_c^a(\beta)+\Delta) = 0$ which, together with (2.17) and $F(\beta,h) \geq 0$, implies the statement. □

5.3 Irrelevance of Disorder for $\alpha < 1/2$ via Replica Coupling

We want to say first of all that, if $0 < \alpha < 1/2$ and β is sufficiently small (i.e., if disorder is sufficiently weak), then $h_c(\beta) = h_c^a(\beta)$. Recalling that $F^a(\beta, h_c^a(\beta)+\Delta) = F(0,\Delta)$, this follows immediately from

Theorem 5.2. *[4, 43] Assume that either* $0 < \alpha < 1/2$ *or that*

$$\alpha = 1/2 \ \text{ and } \ \sum_{n\in\mathbb{N}} n^{-1}L(n)^{-2} < \infty. \tag{5.22}$$

Then, for every $\epsilon > 0$ *there exist* $\beta_0(\epsilon) > 0$ *and* $\Delta_0(\epsilon) > 0$ *such that, for every* $\beta \leq \beta_0(\epsilon)$ *and* $0 < \Delta < \Delta_0(\epsilon)$*, one has*

$$(1-\epsilon)F(0,\Delta) \ \leq \ F(\beta, h_c^a(\beta)+\Delta) \leq F(0,\Delta). \tag{5.23}$$

Observe that this implies in particular that, under the assumptions of the theorem, the exponent ν equals $2 - 1/\alpha$ as in the homogeneous case. Indeed note that, for Δ small,

$$\frac{\log(1-\epsilon) + \log F(0,\Delta)}{\log \Delta} \geq \frac{\log F(\beta, h_c(\beta)+\Delta)}{\log \Delta} \geq \frac{\log F(0,\Delta)}{\log \Delta} \tag{5.24}$$

and the statement follows taking the limit $\Delta \to 0$ from definition (4.5) of the specific heat exponent.

We will see in Section 5.6 that the same cannot hold for $\alpha > 1/2$: in that case, ν is necessarily non-positive in for the quenched system presence of disorder, while it is positive for the annealed system. One

could therefore think that quenched and annealed behaviors are completely different. This is however not completely true. Indeed, the next theorem shows that $F(\beta, h)$ and $F^a(\beta, h)$ are very close, provided that $1/2 \leq \alpha < 1$ if one is not too close to the critical point. More precisely one has

Theorem 5.3. *Assume that* $1/2 < \alpha < 1$. *There exists a slowly varying function* $\check{L}(.)$ *and, for every* $\epsilon > 0$, *constants* $a_1(\epsilon) < \infty$ *and* $\Delta_0(\epsilon) > 0$ *such that, if*

$$a_1(\epsilon)\beta^{2\alpha/(2\alpha-1)}\check{L}(1/\beta) \;\leq\; \Delta \;\leq\; \Delta_0(\epsilon), \tag{5.25}$$

the inequalities (5.23) *hold.*

To see more clearly what this says on the relation between quenched and annealed critical points, forget about the slowly varying functions; then, Theorem 5.3 implies

$$0 \leq h_c(\beta) - h_c^a(\beta) \lesssim \beta^{2\alpha/(2\alpha-1)}.$$

Since $2\alpha/(2\alpha - 1) > 2$, this shows in particular that

$$\lim_{\beta \searrow 0} \frac{h_c(\beta)}{h_c^a(\beta)} = 1. \tag{5.26}$$

Remark 5.4. Theorem 5.3 was proven in [4, Th. 3] and then in [43, Th. 2.2]. The two results differ only in the form of the slowly varying function $\check{L}(.)$. In general, the function $\check{L}(.)$ which pops out from the proof in [43, Th. 2.2] is larger (i.e., worse) than that of [4, Th. 3].

Finally, we consider the "marginal case" $\alpha = \alpha_c = 1/2$ and $\sum_n \frac{1}{nL(n)^2} = \infty$. This is the case, for instance, if $\mathbf{P}$ is the law of the returns of a one-dimensional symmetric random walk, where $L(.)$ is asymptotically constant, as mentioned in Section 2.6. As we mentioned, this case is still debated even in the physical literature. The "most likely" scenario [17] is that disorder is "marginally relevant" in this case: $h_c(\beta) \neq h_c^a(\beta)$ for every positive β, but the two critical points are equal at every order in a weak-disorder perturbation theory. Other works, e.g. [21], claim on the other hand that disorder is irrelevant in this situation.

What one can prove for the moment is the following:

Theorem 5.5. *[4, 43] Assume that* $\alpha = 1/2$ *and* $\sum_{n\in\mathbb{N}} n^{-1}L(n)^{-2} = \infty$. *Let* $\ell(.)$ *be the slowly varying function (diverging at infinity) defined by*

$$\sum_{n=1}^{N} \frac{1}{nL(n)^2} \overset{N\to\infty}{\sim} \ell(N). \tag{5.27}$$

For every $\epsilon > 0$ there exist constants $a_2(\epsilon) < \infty$ and $\Delta_0(\epsilon) > 0$ such that, if $0 < \Delta \leq \Delta_0(\epsilon)$ and if the condition

$$\frac{1}{\beta^2} \geq a_2(\epsilon)\, \ell\left(\frac{a_2(\epsilon)|\log F(0,\Delta)|}{F(0,\Delta)}\right) \tag{5.28}$$

is verified, then Eq. (5.23) *holds.*

Remark 5.6. To be precise, in the statement of [4, Th. 4] the condition (5.28) is replaced by a different one (essentially, the factor $|\log F(0,\Delta)|$ in the argument of $\ell(.)$ does not appear). In this sense, the condition (5.28) under which we prove here (5.23) is not the best possible one. However, for many "reasonable" and physically interesting choices of $L(.)$ in (2.2), Theorem 5.5 and Theorem 4 of [4] are equivalent. In particular, if $\mathbf{P}$ is the law of the returns to zero of the simple random walk $\{S_n\}_{n\geq 0}$ in one dimension, i.e. $\tau = \{n \geq 0 : S_{2n} = 0\}$, in which case $L(.)$ and $\widetilde{L}(.)$ are asymptotically constant and $\ell(N) \sim a_3 \log N$, one sees easily that (5.28) is verified as soon as

$$\Delta \geq a_4(\epsilon) e^{-\frac{a_5(\epsilon)}{\beta^2}}, \tag{5.29}$$

which is the same condition given in [4].

Note, by the way, that in this case the difference $h_c(\beta) - h_c^a(\beta)$ vanishes faster than any power of β, for $\beta \searrow 0$. This confirms the fact that, even if the two critical points can be different, they cannot be distinguished perturbatively.

5.4 Some Open Problems

The results of previous section, while giving rigorous bases to predictions based on the Harris criterion, leave various intriguing gaps in our comprehension of the matter. Let us list a few of them, in random order:

1. Let $\alpha < 1/2$. Does there exist a $\beta_c < \infty$ such that $h_c(\beta) \neq h_c^a(\beta)$ for $\beta > \beta_c$? If yes, how smooth is $h_c(\beta)$ at β_c? Does ν equal $2 - 1/\alpha$ also for β large?
2. Again, let $\alpha < 1/2$ and look at Eq. (5.9). Is it true that

$$F(\beta, h_c^a(\beta) + \Delta) \geq F(0,\Delta) - \frac{\beta^2}{2}(\partial_\Delta F(0,\Delta))^2(1 + O(\beta^2))?$$

3. Under the assumptions of Theorems 5.3 or 5.5, does there exist positive values of β for which quenched and annealed critical points coincide? It is sort of reasonable to conjecture that the answer is "no", at least for $\alpha > 1/2$.

The reader might be tempted to think that such questions should be easy to answer numerically. If so, he should have a look at Ref. [12] where one gets an idea (in the context of random heteropolymers at selective interfaces) of why numerical tests become extremely hard in the neighborhood of the critical curve.

Remark 5.7. Between the time these notes were written and the time they were published, the above open problems have been to a large extent solved. In particular:

- in Ref. [44] it was proven that for every $\alpha > 0$, if β is large enough and, say, ω is Gaussian, then $h_c(\beta) \neq h_c^a(\beta)$.
- The question posed in open problem (2) has been answered positively in Ref. [28], although in a slightly weaker sense.
- In Ref. [16] it was proven that as soon as $\alpha > 1/2$ and $\beta > 0$ one has $h_c(\beta) \neq h_c^a(\beta)$.

5.5 Proof of Theorems 5.2–5.5

We follow the approach of [43] which, with respect to that of [4], has the advantage of technical simplicity and of being closely related to the interpolation ideas of Section 5.2. On the other hand, we encourage the reader to look also at the methods developed in [4], which have the bonus of extending in a natural way beyond the Gaussian case and of giving in some cases sharper results (cf. Remarks 5.4 and 5.6 above).

A natural idea to show that quenched and annealed systems have (approximately) the same free energy is to apply the *second moment method*: one computes $\mathbb{E}(Z_N(\beta, h))$ and $\mathbb{E}((Z_N(\beta, h))^2)$ and if it happens that the ratio

$$\frac{[\mathbb{E}Z_{N,\omega}(\beta, h)]^2}{\mathbb{E}[(Z_{N,\omega}(\beta, h))^2]} \tag{5.30}$$

remains positive for $N \to \infty$, or at least it vanishes slower than exponentially, it is not difficult to deduce that $F(\beta, h) = F^a(\beta, h)$. This approach has turned out to be very powerful for instance in controlling the high-temperature phase of the Sherrington-Kirkpatrick mean field model in absence of magnetic field [40, Ch. 2.2]. However, this

simple idea does not work in our case and the ratio (5.30) vanishes exponentially for every $\beta, \Delta > 0$. This is not surprising after all, since we already know from Theorem 5.1 that quenched and annealed free energy *do not* coincide. There are two possible ways out of this problem. One is to perform the second moment method not on the system of size N but on a smaller system whose size $N(\Delta)$ remains finite as long as Δ is positive and fixed, and diverges only for $\Delta \to 0$. If $N(\Delta)$ is chosen to be the *correlation length of the annealed system*, one can see that on this scale the ratio (5.30) stays positive, so that $F_{N(\Delta)}(\beta, h_c^a(\beta) + \Delta) \simeq F_{N(\Delta)}(0, \Delta)$. One is then left with the delicate problem of glueing together many blocks of size $N(\Delta)$ to obtain an estimate of the type $F(\beta, h_c^a(\beta) + \Delta) \geq (1 - \epsilon)F(0, \Delta)$ for the full free energy. This is, in very rough words, the approach of Ref. [4]. The other possibility, which we are going to present, is to abandon the second moment idea in favor of a generalization of the *replica coupling method* [31] [43]. This method was introduced in [31] in the context of mean field spin glasses and gives a very efficient control of the Sherrington-Kirkpatrick model at high temperature (β small), i.e., for weak disorder, which is the same situation we are after here.

The two methods are in reality not orthogonal: they share the idea that the important object to look at is the intersection of two independent renewals $\tau^{(1)}, \tau^{(2)}$. To see why this quantity arises naturally, let us compute the second moment of the partition function. If $\tau^{(1)}, \tau^{(2)}$ are independent renewal processes with product law $\mathbf{P}^{\otimes 2}(.)$, recalling the definition $\Delta = h + \beta^2/2$, one can write

$$\begin{aligned}
&\mathbb{E}((Z_{N,\omega}(\beta,h))^2) = \\
&\qquad \mathbb{E}\,\mathbf{E}^{\otimes 2}\left(e^{\sum_{n=1}^N (\beta\omega_n + h)(\mathbf{1}_{\{n\in\tau^{(1)}\}} + \mathbf{1}_{\{n\in\tau^{(2)}\}})} \mathbf{1}_{\{N\in\tau^{(1)}\}}\mathbf{1}_{\{N\in\tau^{(2)}\}}\right) = \\
&= \mathbf{E}^{\otimes 2}\big[e^{\Delta(|\tau^{(1)}\cap\{1,\dots,N\}| + |\tau^{(2)}\cap\{1,\dots,N\}|) + \beta^2|\tau^{(1)}\cap\tau^{(2)}\cap\{1,\dots,N\}|}\mathbf{1}_{\{N\in\tau^{(1)}\}}\mathbf{1}_{\{N\in\tau^{(2)}\}}\big].
\end{aligned} \tag{5.31}$$

Considering also that

$$[\mathbb{E}\, Z_{N,\omega}(\beta,h)]^2 = \mathbf{E}^{\otimes 2}\left(e^{\Delta|(\tau^{(1)}\cap\{1,\dots,N\}| + |\tau^{(2)}\cap\{1,\dots,N\}|)}\mathbf{1}_{\{N\in\tau^{(1)}\}}\mathbf{1}_{\{N\in\tau^{(2)}\}}\right)$$

one sees that the ratio (5.30) depends on the typical number of points that $\tau^{(1)}$ and $\tau^{(2)}$ have in common up to time N. One sees also why this ratio has to vanish exponentially $N \to \infty$: as long as $\Delta > 0$ the renewals $\tau^{(i)}$, with law modified by the factor $\exp(\Delta|\tau^{(1)} \cap \{1, \dots, N\}|)$, are finitely recurrent and therefore will have a number of intersections in $\{1, \dots, N\}$ which grows proportionally to N.

Proof of Theorem 5.2. The second inequality in (5.23) is just Eq. (2.17). As for the first one, let $\Delta > 0$ and recall identity (5.2). Define, in analogy with (5.10),

$$R_{N,\Delta}(t,\beta) := \frac{1}{N}\mathbb{E}\log\left\langle e^{\sum_{n=1}^{N}(\beta\sqrt{t}\omega_n - t\beta^2/2)\delta_n}\right\rangle_{\Delta,N} \tag{5.32}$$

for $0 \le t \le 1$ (to the purpose of Theorem 5.2 we do not need the variational parameter q) where the measure $\langle . \rangle_{N,\Delta}$ was defined after Eq. (5.1). Observe that

$$R_{N,\Delta}(0,\beta) = 0 \tag{5.33}$$

while

$$R_{N,\Delta}(1,\beta) = R_{N,\Delta}(\beta). \tag{5.34}$$

As for the t-derivative one finds (just take (5.17) and put $q = 0$):

$$\frac{\mathrm{d}}{\mathrm{d}t}R_{N,\Delta}(t,\beta) = -\frac{\beta^2}{2N}\sum_{m=1}^{N}\mathbb{E}\left\{\left(\frac{\left\langle \delta_m\, e^{\sum_{n=1}^{N}(\beta\sqrt{t}\omega_n - t\beta^2/2)\delta_n}\right\rangle_{\Delta,N}}{\left\langle e^{\sum_{n=1}^{N}(\beta\sqrt{t}\omega_n - t\beta^2/2)\delta_n}\right\rangle_{\Delta,N}}\right)^2\right\}. \tag{5.35}$$

Recall definition (5.13) (specialized to the case $q = 0$) of the random measure $\langle . \rangle_{N,\Delta,t}$ and let $\langle . \rangle_{N,\Delta,t}^{\otimes 2}$ be the product measure acting on the pair $(\tau^{(1)}, \tau^{(2)})$, while $\delta_n^{(i)} := \mathbf{1}_{\{n\in\tau^{(i)}\}}$. Note that the two *replicas* $\tau^{(i)}, i = 1, 2$ are subject to the *same realization* ω *of disorder*. Then, one can rewrite

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}R_{N,\Delta}(t,\beta) &= -\frac{\beta^2}{2N}\mathbb{E}\sum_{m=1}^{N}\left\langle \delta_m^{(1)}\delta_m^{(2)}\right\rangle_{N,\Delta,t}^{\otimes 2} = \\ &= -\frac{\beta^2}{2N}\mathbb{E}\left\langle \left|\tau^{(1)}\cap\tau^{(2)}\cap\{1,\dots,N\}\right|\right\rangle_{N,\Delta,t}^{\otimes 2}.\end{aligned} \tag{5.36}$$

Since we need a lower bound for $R_{N,\Delta}(\beta)$ to prove the first inequality in (5.23), the fact that this derivative is non-positive seems to go in the wrong direction. Let us not lose faith and let us define, for $\lambda \ge 0$,

$$\begin{aligned}R_{N,\Delta}^{(2)}(t,\lambda,\beta) &:= \frac{1}{2N}\mathbb{E}\log\left\langle e^{H_N(t,\lambda,\beta;\tau^{(1)},\tau^{(2)})}\right\rangle_{N,\Delta}^{\otimes 2} \\ &:= \frac{1}{2N}\mathbb{E}\log\left\langle e^{\sum_{n=1}^{N}(\beta\sqrt{t}\omega_n - t\beta^2/2)(\delta_n^{(1)}+\delta_n^{(2)})+\lambda\beta^2\sum_{n=1}^{N}\delta_n^{(1)}\delta_n^{(2)}}\right\rangle_{N,\Delta}^{\otimes 2}\end{aligned} \tag{5.37}$$

where the product measure $\langle . \rangle_{N,\Delta}^{\otimes 2}$ acts on the pair $(\tau^{(1)}, \tau^{(2)})$. The index "$^{(2)}$" refers to the fact that this quantity involves two copies (*replicas*) of the system. Observe that we are letting the two replicas interact through a term which is positive, extensive (i.e., of order N) and closely related to what appears in the right-hand side of Eq. (5.36). Note also that

$$R^{(2)}_{N,\Delta}(0,\lambda,\beta) = \frac{1}{2N} \log \left\langle e^{\lambda\beta^2 \sum_{n=1}^N \delta_n^{(1)}\delta_n^{(2)}} \right\rangle_{N,\Delta}^{\otimes 2}, \tag{5.38}$$

while the factor 2 in the denominator guarantees that

$$R^{(2)}_{N,\Delta}(t,0,\beta) = R_{N,\Delta}(t,\beta). \tag{5.39}$$

Again via integration by parts (the computation is conceptually as easy as the one which led to Eq. (5.17)),

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} R^{(2)}_{N,\Delta}(t,\lambda,\beta) &= \frac{\beta^2}{2N} \sum_{m=1}^N \mathbb{E} \frac{\left\langle \delta_m^{(1)}\delta_m^{(2)} e^{H_N(t,\lambda,\beta;\tau^{(1)},\tau^{(2)})} \right\rangle_{N,\Delta}^{\otimes 2}}{\left\langle e^{H_N(t,\lambda,\beta;\tau^{(1)},\tau^{(2)})} \right\rangle_{N,\Delta}^{\otimes 2}} - \\
&- \frac{\beta^2}{4N} \sum_{m=1}^N \mathbb{E} \left\{ \left(\frac{\left\langle (\delta_m^{(1)} + \delta_m^{(2)}) e^{H_N(t,\lambda,\beta;\tau^{(1)},\tau^{(2)})} \right\rangle_{N,\Delta}^{\otimes 2}}{\left\langle e^{H_N(t,\lambda,\beta;\tau^{(1)},\tau^{(2)})} \right\rangle_{N,\Delta}^{\otimes 2}} \right)^2 \right\} \leq \\
&\leq \frac{\beta^2}{2N} \mathbb{E} \sum_{m=1}^N \frac{\left\langle \delta_m^{(1)}\delta_m^{(2)} e^{H_N(t,\lambda,\beta;\tau^{(1)},\tau^{(2)})} \right\rangle_{N,\Delta}^{\otimes 2}}{\left\langle e^{H_N(t,\lambda,\beta;\tau^{(1)},\tau^{(2)})} \right\rangle_{N,\Delta}^{\otimes 2}} = \frac{\mathrm{d}}{\mathrm{d}\lambda} R^{(2)}_{N,\Delta}(t,\lambda,\beta).
\end{aligned} \tag{5.40}$$

This can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t} R^{(2)}_{N,\Delta}(t,\lambda - t,\beta) \leq 0$$

which implies that, for every $0 \leq t \leq 1$ and λ,

$$R^{(2)}_{N,\Delta}(t,\lambda,\beta) \leq R^{(2)}_{N,\Delta}(0,\lambda + t,\beta). \tag{5.41}$$

Going back to Eqs. (5.35) and the last equality in (5.40) and using the fact that for every convex function $\psi(.)$ one has $x\psi'(0) \leq \psi(x) - \psi(0)$ one finds

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(-R_{N,\Delta}(t,\beta)\right) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\, R^{(2)}_{N,\Delta}(t,\lambda,\beta)\Big|_{\lambda=0} \leq$$
$$\leq \frac{R^{(2)}_{N,\Delta}(t,2-t,\beta) - R^{(2)}_{N,\Delta}(t,0,\beta)}{2-t}. \quad (5.42)$$

Finally, using monotonicity of $R^{(2)}_{N,\Delta}(t,\lambda,\beta)$ with respect to λ and (5.39), one obtains the bound

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(-R_{N,\Delta}(t,\beta)\right) \leq R^{(2)}_{N,\Delta}(0,2,\beta) + \left(-R_{N,\Delta}(t,\beta)\right), \quad (5.43)$$

where we used (5.41) and the fact that $2-t \geq 1$ (of course, we could have chosen $1+\eta-t$ instead of $2-t$ for some $\eta > 0$ in (5.42) and the estimates would be modified in a straightforward way). We can now integrate with respect to t between 0 and 1 this differential inequality (or use Gronwall's Lemma, if you prefer) and, recalling Eqs. (5.34) and (5.33), we obtain

$$-(e-1)R^{(2)}_{N,\Delta}(0,2,\beta) \leq R_{N,\Delta}(\beta) \leq 0. \quad (5.44)$$

Before we proceed, we would like to summarize what we did so far. To prove Theorem 5.2 we need the *lower bound* $\lim_{N\to\infty} R_{N,\Delta}(\beta) \geq -\epsilon F(0,\Delta)$ but, as in Section 5.2, it seems that the interpolation method gives rather *upper bounds* on $R_{N,\Delta}(\beta)$. Then, through the replica coupling trick we transferred this problem into the problem of proving an *upper bound* for a quantity, $R^{(2)}_{N,\Delta}(t,\lambda,\beta)$, which is analogous to $R_{N,\Delta}(\beta)$, except that it involves two interacting copies of the system. Moreover, by throwing away a (complicated, but with a definite sign) term in Eq. (5.40), we reduced to the problem of bounding from above $R^{(2)}_{N,\Delta}(t,\lambda,\beta)$ computed at $t=0$. In other words, we replaced the task of estimating from below $R_{N,\Delta}(\beta)$ with that of estimating from above a quantity which involves no quenched disorder, and which is therefore easier to analyze. While this procedure might look a bit magic, the basic underlying idea is the following. $R^{(2)}_{N,\Delta}(t,\lambda,\beta)$ is obviously non-decreasing as a function of λ. Suppose however that, for some $\lambda > 0$, $R^{(2)}_{N,\Delta}(t,\lambda,\beta)$ is not very different from the value it has at $\lambda = 0$ (of course, proving this amounts to proving an *upper* bound on $R^{(2)}_{N,\Delta}(t,\lambda,\beta)$.) Then, looking at the definition (5.37), this means that the cardinality of the intersection $\tau^{(1)} \cap \tau^{(2)} \cap \{1,\ldots,N\}$ is typically not large and this, through Eqs. (5.33), (5.34) and (5.36) implies a *lower* bound on $R_{N,\Delta}(\beta)$.

Let us now restart from (5.44) and note that

$$\begin{aligned} &R^{(2)}_{N,\Delta}(0,2,\beta) \;= \\ &= -F_N(0,\Delta)+\frac{1}{2N}\log \mathbf{E}^{\otimes 2}\Big(e^{2\beta^2\sum_{n=1}^N \delta_n^{(1)}\delta_n^{(2)}+\Delta\sum_{n=1}^N(\delta_n^{(1)}+\delta_n^{(2)})\delta_N^{(1)}\delta_N^{(2)}}\Big)\leq \\ &\leq\; -F_N(0,\Delta)+\frac{F_N(0,q\Delta)}{q}+\frac{1}{2Np}\log \mathbf{E}^{\otimes 2}\left(e^{2p\beta^2\sum_{n=1}^N \delta_n^{(1)}\delta_n^{(2)}}\right) \end{aligned} \tag{5.45}$$

where we used Hölder's inequality and the positive numbers p and q (satisfying $1/p+1/q=1$) are to be determined. Taking the thermodynamic limit,

$$\begin{aligned} \limsup_{N\to\infty} R^{(2)}_{N,\Delta}(0,2,\beta) \leq \limsup_{N\to\infty}\frac{1}{2Np}\log \mathbf{E}^{\otimes 2}\left(e^{2p\beta^2\sum_{n=1}^N \delta_n^{(1)}\delta_n^{(2)}}\right) \\ + F(0,\Delta)\left(\frac{1}{q}\frac{F(0,q\Delta)}{F(0,\Delta)}-1\right). \end{aligned} \tag{5.46}$$

But we know from the expression (4.2) of the free energy of the homogeneous system and from the property (2.3) of slow variation that, for every $q>0$,

$$\lim_{\Delta\to 0^+}\frac{F(0,q\Delta)}{F(0,\Delta)} = q^{1/\alpha}. \tag{5.47}$$

Therefore, taking $q=q(\epsilon)$ sufficiently close to (but strictly larger than) 1 and $\Delta_0(\epsilon)>0$ sufficiently small one has, uniformly on $\beta\geq 0$ and on $0<\Delta\leq\Delta_0(\epsilon)$,

$$\begin{aligned} &\limsup_{N\to\infty} R^{(2)}_{N,\Delta}(0,2,\beta) \leq \\ &\leq \frac{\epsilon}{e-1}F(0,\Delta)+\limsup_{N\to\infty}\frac{1}{2Np(\epsilon)}\log \mathbf{E}^{\otimes 2}\left(e^{2p(\epsilon)\beta^2\sum_{n=1}^N \delta_n^{(1)}\delta_n^{(2)}}\right). \end{aligned} \tag{5.48}$$

Of course, $p(\epsilon):=q(\epsilon)/(q(\epsilon)-1)<\infty$ as long as $\epsilon>0$. Note that, in view of (5.44), Theorem 5.2 would be proved if the second term in the right-hand side of (5.48) were zero. Up to now, we have not used yet the assumption that $\alpha<1/2$ or that (5.22) holds, but now the right moment has come. The way this assumption enters the game is that it guarantees that the renewal $\tau^{(1)}\cap\tau^{(2)}$ is transient under the law $\mathbf{P}^{\otimes 2}$. Indeed,

$$\mathbf{E}^{\otimes 2}\left(\sum_{n\geq 1}\mathbf{1}_{n\in\tau^{(1)}\cap\tau^{(2)}}\right)=\sum_{n\geq 1}\mathbf{P}(n\in\tau)^2<\infty \tag{5.49}$$

since, as proven in [18],

$$\mathbf{P}(n\in\tau)\overset{n\to\infty}{\sim}\frac{C_\alpha}{L(n)n^{1-\alpha}}:=\frac{\alpha\sin(\pi\alpha)}{\pi}\frac{1}{L(n)n^{1-\alpha}}. \tag{5.50}$$

Actually, Eq. (5.50) holds more generally for $0<\alpha<1$ and we will need it to prove Theorems 5.3 and 5.5.

Transience and renewal properties of the process of $\tau^{(1)}\cap\tau^{(2)}$ implies that

$$\mathbf{P}^{\otimes 2}(|\tau^{(1)}\cap\tau^{(2)}|\geq k)\leq(1-c)^k, \tag{5.51}$$

for some $0<c<1$: after each "renewal epoch", i.e., each point of $\tau^{(1)}\cap\tau^{(2)}$, the intersection renewal has a positive probability c of jumping to infinity. Therefore, there exists $\beta_1>0$ such that

$$\sup_N \mathbf{E}^{\otimes 2}\left(e^{2p(\epsilon)\beta^2\sum_{n=1}^N\delta_n^{(1)}\delta_n^{(2)}}\right)<\infty \tag{5.52}$$

for every $\beta^2 p(\epsilon)\leq\beta_1^2$. Together with (5.48) and (5.2), this implies

$$F(\beta,h_c^a(\beta)+\Delta)\geq(1-\epsilon)F(0,\Delta) \tag{5.53}$$

as soon as $\beta^2\leq\beta_0^2(\epsilon):=\beta_1^2/p(\epsilon)$, and therefore the validity of Theorem 5.2. □

Proof of Theorem 5.3. In what follows we assume that Δ is sufficiently small so that $F(0,\Delta)\ll 1$. For simplicity of exposition, we assume also that $L(.)$ tends to a positive constant $L(\infty)$ at infinity (for the general case, which is not significantly more difficult, cf. [43]).

If we try to repeat the proof of Theorem 5.3 in this case, what goes wrong is that the intersection $\tau^{(1)}\cap\tau^{(2)}$ is now recurrent, so that (5.52) does not hold any more. The natural idea is then not to let N tend to infinity at Δ fixed, but rather to work on a system of size $N(\Delta)$, which diverges only when $\Delta\to 0$, i.e., when the annealed critical point is approached. In particular, we let $N=N(\Delta):=c|\log F(0,\Delta)|/F(0,\Delta)$ with $c>0$ large to be fixed later. Note also that this choice of $N(\Delta)$ is quite similar to that made in [4], where one applies the second moment method on a system of size $c/F(0,\Delta)$ with c large. This choice has a clear physical meaning: indeed, we will see in Section 6 that the correlation functions of the annealed system decay exponentially on

distances of order $1/F(0,\Delta)$ (the logarithmic factor in our definition of $N(\Delta)$ should be seen as a technical necessity).

By the superadditivity property (2.11) we have, in analogy with (5.1),

$$F(\beta, -\beta^2/2 + \Delta) \geq F_{N(\Delta)}(0,\Delta) + R_{N(\Delta),\Delta}(\beta). \tag{5.54}$$

To prove Theorem 5.3 we need to show that the first term in the right-hand side of (5.54) is essentially $F(0,\Delta)$, while the second is not smaller than $-\epsilon F(0,\Delta)$, in the range of parameters determined by condition (5.25). The first fact is easy: as follows from Proposition 2.7 of [25], there exists $a_6 \in (0,\infty)$ (depending only on the law $K(.)$ of the renewal) such that

$$F_N(0,\Delta) \geq F(0,\Delta) - a_6 \frac{\log N}{N} \tag{5.55}$$

for every N. Choosing $c = c(\epsilon)$ large enough, Eq. (5.55) implies that

$$F_{N(\Delta)}(0,\Delta) \geq (1-\epsilon) F(0,\Delta). \tag{5.56}$$

As for $R_{N(\Delta),\Delta}(\beta)$, we have from Eqs. (5.44) and (5.45)

$$\begin{aligned} \frac{R_{N(\Delta),\Delta}(\beta)}{e-1} \;\geq\; & -F(0,\Delta)\left(\frac{1}{q}\frac{F(0,q\Delta)}{F(0,\Delta)} - 1\right) - \epsilon F(0,\Delta) \\ & - \frac{1}{2N(\Delta)p} \log \mathbf{E}^{\otimes 2}\left(e^{2p\beta^2 \sum_{n=1}^{N(\Delta)} \delta_n^{(1)}\delta_n^{(2)}}\right), \end{aligned} \tag{5.57}$$

where we used Eqs. (5.56) and (2.11) to bound $-(1/q)F_{N(\Delta)}(0,q\Delta) + F_{N(\Delta)}(0,\Delta)$ from below. Choosing again $q = q(\epsilon)$ we obtain, for $\Delta \leq \Delta_0(\epsilon)$,

$$\frac{R_{N(\Delta),\Delta}(\beta)}{e-1} \geq -2\epsilon F(0,\Delta) - \frac{1}{2N(\Delta)p(\epsilon)} \log \mathbf{E}^{\otimes 2}\left(e^{2p(\epsilon)\beta^2 \sum_{n=1}^{N(\Delta)} \delta_n^{(1)}\delta_n^{(2)}}\right). \tag{5.58}$$

It was proven in [4, Lemma 3] and [43, Section 3.1] that if $1/2 < \alpha < 1$ there exists $a_7 = \in (0,\infty)$, which depends in particular on $L(\infty)$, such that for every integers N and k

$$\mathbf{P}^{\otimes 2}\left(\left|\tau^{(1)} \cap \tau^{(2)} \cap \{1,\ldots,N\}\right| \geq k\right) \leq \left(1 - \frac{a_7}{N^{2\alpha-1}}\right)^k, \tag{5.59}$$

which should be compared with (5.51), valid for $\alpha < 1/2$. Thanks to the geometric bound (5.59) we have

$$\mathbf{E}^{\otimes 2}\left(e^{2p(\epsilon)\beta^2\sum_{n=1}^{N(\Delta)}\delta_n^{(1)}\delta_n^{(2)}}\right)=\sum_{k\geq 0}\mathbf{P}^{\otimes 2}\left(\sum_{n=1}^{N(\Delta)}\delta_n^{(1)}\delta_n^{(2)}=k\right)e^{2p(\epsilon)\beta^2 k}$$
$$\leq\left(1-e^{2\beta^2 p(\epsilon)}\left(1-\frac{a_7}{N(\Delta)^{2\alpha-1}}\right)\right)^{-1},\quad(5.60)$$

whenever

$$e^{2\beta^2 p(\epsilon)}\left(1-\frac{a_7}{N(\Delta)^{2\alpha-1}}\right)<1$$

and this is of course the case if

$$e^{2\beta^2 p(\epsilon)}\left(1-\frac{a_7}{N(\Delta)^{2\alpha-1}}\right)\leq\left(1-\frac{a_7}{2N(\Delta)^{2\alpha-1}}\right).\quad(5.61)$$

At this point, using the definition of $N(\Delta)$ and point (2) of Theorem 4.1, it is not difficult to see that there exists a positive constant $a_8(\epsilon)$ such that (5.61) holds if

$$\beta^2 p(\epsilon)\leq a_8(\epsilon)\frac{\Delta^{(2\alpha-1)/\alpha}}{|\log F(0,\Delta)|^{2\alpha-1}}.\quad(5.62)$$

Condition (5.62) is equivalent to the first inequality in (5.25), for a suitable choice of $a_1(\epsilon)$ and $\check{L}(.)$. As a consequence, for $N(\Delta)$ sufficiently large (i.e., for Δ sufficiently small)

$$\frac{1}{2N(\Delta)p(\epsilon)}\log\mathbf{E}^{\otimes 2}\left(e^{2p(\epsilon)\beta^2\sum_{n=1}^{N(\Delta)}\delta_n^{(1)}\delta_n^{(2)}}\right)\leq$$
$$\leq\frac{F(0,\Delta)}{2c(\epsilon)p(\epsilon)|\log F(0,\Delta)|}\log\left(\frac{2N(\Delta)^{2\alpha-1}}{a_7}\right).\quad(5.63)$$

Recalling Eq. (4.2) one sees that, if $c(\epsilon)$ is chosen large enough,

$$\frac{1}{2N(\Delta)p(\epsilon)}\log\mathbf{E}^{\otimes 2}\left(e^{2p(\epsilon)\beta^2\sum_{n=1}^{N(\Delta)}\delta_n^{(1)}\delta_n^{(2)}}\right)\leq\epsilon F(0,\Delta).\quad(5.64)$$

Together with Eqs. (5.54), (5.56) and (5.58), this concludes the proof of the theorem. □

Proof of Theorem 5.5. The proof is almost identical to that of Theorem 5.3 and up to Eq. (5.58) no changes are needed. One has however to be careful with the geometric bound (5.59): in this case, it is not sufficient to replace α by $1/2$, since the behavior at infinity of the slowly varying function $L(.)$ in (2.2) is here essential. The correct bound in this case is (cf. [4, Lemma 3] and [43, Sec. 3.1])

$$\mathbf{P}^{\otimes 2}\left(\sum_{n=1}^{N}\delta_n^{(1)}\delta_n^{(2)}\geq k\right)\leq\left(1-\frac{a_9}{\ell(N)}\right)^k. \tag{5.65}$$

for every N, for some $a_9>0$. We recall that $\ell(.)$ is the slowly varying function, diverging at infinity, defined by (5.27). In analogy with Eq. (5.60) one obtains then

$$\mathbf{E}^{\otimes 2}\left(e^{2p(\epsilon)\beta^2\sum_{n=1}^{N(\Delta)}\delta_n^{(1)}\delta_n^{(2)}}\right)\leq\left(1-e^{2\beta^2p(\epsilon)}\left(1-\frac{a_9}{\ell(N(\Delta))}\right)\right)^{-1} \tag{5.66}$$

whenever the right-hand side is positive. Choosing $a_2(\epsilon)$ large enough one sees that if condition (5.28) is fulfilled then

$$e^{2\beta^2p(\epsilon)}\left(1-\frac{a_9}{\ell(N(\Delta))}\right)\leq\left(1-\frac{a_9}{2\ell(N(\Delta))}\right) \tag{5.67}$$

and, in analogy with (5.63),

$$\frac{1}{2N(\Delta)p(\epsilon)}\log\mathbf{E}^{\otimes 2}\left(e^{2(\epsilon)\beta^2\sum_{n=1}^{N(\Delta)}\delta_n^{(1)}\delta_n^{(2)}}\right)\leq$$
$$\leq\frac{F(0,\Delta)}{2c(\epsilon)p(\epsilon)|\log F(0,\Delta)|}\log\left(\frac{2\ell(N(\Delta))}{a_9}\right). \tag{5.68}$$

From this estimate, for $c(\epsilon)$ sufficiently large one obtains again (5.64) and as a consequence the statement of Theorem 5.5. □

5.6 Smoothing Effect of Disorder (Relevance for $\alpha > 1/2$)

Section 5.3 was devoted to showing that, for $\alpha<\alpha_c$, (weak) disorder is irrelevant, in that it does not change the specific heat exponent ν and in that the transition point coincides with the annealed one as long as β is small. We saw also that for $\alpha_c\leq\alpha<1$ quenched and annealed free energies and critical points are very close (Theorems 5.3 and 5.5). This might leave the reader with the doubt that disorder might be irrelevant in this situation too. The purpose of the present section is to show that this is not the case.

We start by recalling that via Theorem 4.1 and (2.17) we know that $F(\beta,h_c^a(\beta)+\Delta)\lesssim\Delta^{\max(1/\alpha,1)}$. This bound is however quite poor: if we go back to (5.8) and we choose $q=\Delta/\beta^2$ we obtain

$$F(\beta,h_c^a(\beta)+\Delta)\leq\frac{\Delta^2}{2\beta^2} \tag{5.69}$$

which is better, for Δ small and $\alpha > 1/2$. The point is however that, since one expects that $h_c^a(\beta) \neq h_c(\beta)$ in this situation, (5.69) does not say anything about the critical behavior of the quenched system; for this, we would need rather an upper bound on $F(\beta, h_c(\beta) + \Delta)$. This is just the content of the following result, which we state in the case of Gaussian disorder:

Theorem 5.8. *[26, 27] For every $\beta > 0$, $\alpha > 0$ and $\Delta > 0$ one has*

$$F(\beta, h_c(\beta) + \Delta) \leq \frac{(1+\alpha)}{2\beta^2} \Delta^2. \tag{5.70}$$

Remark 5.9. Theorem 5.8 actually holds beyond the Gaussian case; for instance, in the case of bounded variables ω_n. In this case the statement has to be modified in that the factor 2 in that the denominator in the right-hand side of (5.70) is replaced by $c := c(\mathbb{P})$, a constant which depends only on the disorder distribution $\mathbb{P}$, and the results holds only provided Δ is sufficiently small: $\Delta \leq \Delta_0(\mathbb{P})$, see [26].

Remark 5.10. An obvious implication of Theorem 5.8 is that $\nu \leq 0$ as soon as $\beta > 0$. In this sense, this result is much reminiscent of what was proven in [13, 14] about the specific heat exponent for the nearest-neighbor disordered Ising ferromagnet.

In particular, Theorem 5.8 shows that the specific heat exponent is modified by an arbitrary amount of disorder if $\alpha > \alpha_c$: the phase transition is smoothed by randomness if $\alpha > \alpha_c$ and becomes at least of second order (the effect is particularly dramatic for $\alpha > 1$, where the transition is of first order for $\beta = 0$).

It is also interesting to compare Theorem 5.8 with the celebrated result by M. Aizenman and J. Wehr [2] which states that first order phase transition in spin systems with discrete spin-flip symmetry are smoothed by disorder as long as the spatial dimension verifies $d \leq 2$, while the same holds for $d \leq 4$ if the symmetry is continuous.

A less obvious consequence of Theorem 5.8 is the following:

Theorem 5.11. *[41] Let $\beta > 0$ and $0 \leq \alpha < \infty$. There exists $c > 0$ such that*

$$\lim_{N\to\infty} \mathbb{E}\, \mathbf{P}_{N,\omega}^{\beta,h_c(\beta)} \Big(|\tau \cap \{1, \ldots, N\}| \geq cN^{2/3} \log N \Big) = 0. \tag{5.71}$$

Moreover, under the assumptions of Theorem 5.2, for β sufficiently small

$$\lim_{N\to\infty} \mathbb{E}\, \mathbf{P}_{N,\omega}^{\beta,h_c(\beta)} \Big(|\tau \cap \{1, \ldots, N\}| \geq cN^{2\alpha/(1+\alpha)} \log N \Big) = 0. \tag{5.72}$$

This result should be read as follows. The fact that the transition is at least of second order in presence of disorder implies already that the Gibbs average of the contact fraction defined by (2.14) tends to zero in the thermodynamic limit at the critical point. The additional information provided by Theorem 5.11 are finite-N estimates on the size of $\tau \cap \{1, \dots, N\}$ at criticality. Whether the exponent 2/3 in Eq. (5.71) is optimal or not is an intriguing open question.

Theorem 5.11 was proven in [41][3] (together with more refined finite-size estimates on $\mathbf{E}^{\beta,h}_{N,\omega}(|\tau \cap \{1, \dots, N\}|)$ for $h - h_c(\beta)$ going to zero with N), apart from Eq. (5.72) which is a consequence of [41, Th. 3.1] plus Theorem 5.2 (cf. also Remark 3.2 in [41]).

Proof of Theorem 5.8 (sketch) For a fully detailed proof we refer to [26]. In the case of Gaussian disorder a simpler proof is hinted at in [27] and fully developed in [22, Section 5.4].

Here we give just a sketchy idea of why the transition cannot be of first order when $\beta > 0$. Assume by contradiction that

$$F(\beta, h_c(\beta) + \Delta) \sim c\Delta \quad \text{for } \Delta \to 0^+, \tag{5.73}$$

and consider the system at the critical point $(\beta, h_c(\beta))$. Divide the system of size N into N/M blocks B_i of size M, with the idea that $1 \ll M \ll N$. For a given realization of ω mark the blocks where the empirical average of ω, i.e., $(1/M)\sum_{n \in B_i} \omega_n$ equals approximately Δ/β. By standard large deviation estimates, there are typically $\mathcal{N}_{marked} := (N/M)e^{-M\Delta^2/(2\beta^2)}$ such blocks, the typical distance between two successive ones being $D_{typ} := Me^{M\Delta^2/(2\beta^2)}$. It is a standard fact that if we take M IID standard Gaussian variables and we condition on their empirical average to be δ, for M large they are (roughly speaking) distributed like IID Gaussian variables of variance 1 *and average* δ. Therefore, in a marked block the system sees effective thermodynamic parameters $(\beta_{eff}, h_{eff}) := (\beta, h_c(\beta) + \Delta)$. Now we want to show that the assumption (5.73) leads to the (obviously false) conclusion that $F(\beta, h_c(\beta)) > 0$. Indeed, let $\mathcal{S}_\omega$ be the set of τ configurations such that:

- there are no points of τ in unmarked blocks
- the boundaries of all marked blocks belong to τ.

[3] Theorem 3.1 in [41] is formulated in the case of bounded random variables ω_n, but it generalizes immediately to the Gaussian because the basic ingredient one needs is the concentration inequality [41, Eq. (5.2)], which holds in the case of Gaussian randomness as well.

Note that $\mathcal{S}_\omega$ depends on disorder through the location and the number of marked blocks, and that there is no restriction on τ inside marked blocks. One has the obvious bound

$$F_N(\beta, h_c(\beta)) \geq \frac{1}{N}\mathbb{E}\log \mathbf{E}\Big(e^{\sum_{n=1}^N (\beta\omega_n+h)\delta_n}\mathbf{1}_{\{\tau\in\mathcal{S}_\omega\}}\delta_N\Big). \quad (5.74)$$

But due to the definition of the set $\mathcal{S}_\omega$, the restricted free energy in the right-hand side of (5.74) gets (for M large) a contribution $\mathcal{N}_{marked}\times(M/N)F(\beta, h_c(\beta)+\Delta)$ from marked blocks, and an entropic term $\mathcal{N}_{marked}/N \times \log K(D_{typ})$ from the excursions between marked blocks. Summing the two contributions, recalling the asymptotic behavior (2.2) of $K(.)$, the expression of $\mathcal{N}_{marked}$ and D_{typ} and taking the $N\to\infty$ limit *at M large but fixed* one obtains then

$$F(\beta, h_c(\beta)+\Delta) \geq e^{-M\Delta^2/(2\beta^2)}\left(F(\beta, h_c(\beta)+\Delta) - (1+\alpha)\frac{\Delta^2}{2\beta^2}\right). \quad (5.75)$$

Since the left-hand side of (5.75) is zero, for Δ small and $\beta > 0$ this inequality is clearly in contradiction with the assumption (5.73) that the transition if of first order (actually, even with the assumption $F(\beta, h_c(\beta)+\Delta) \sim c\Delta^y$ with $y < 2$).

6 Correlation Lengths and their Critical Behavior

From certain points of view, the localized region $\mathcal{L}$ is analogous to the high-temperature phase of a spin system. Indeed, in this region one can prove typical high-temperature results like the following: free energy fluctuations are Gaussian on the scale $1/\sqrt{N}$ [3, 25], the infinite-volume Gibbs measure is almost-surely unique and ergodic [9], the free energy is infinitely differentiable, finite-size corrections to the infinite volume free energy are of order $O(1/N)$, and truncated correlation functions decay exponentially with distance [25]. In this section we concentrate on the last point, which turns out to be more subtle than expected, in particular when one approaches the critical line.

In this section we assume that the random variables ω_n are bounded, because the results we mention have been proved in the literature under this assumption. They should however reasonably extend to more general situations, for instance to the Gaussian case.

In the following, $\mathbf{P}^{\beta,h}_{\infty,\omega}(.)$ will denote the infinite-volume Gibbs measure, defined as follows: first of all we modify definitions (2.6) and (2.8) replacing $\sum_{n=1}^N(\beta\omega_n+h)\delta_n$ by

$$\sum_{n=-\lfloor N/2\rfloor}^{\lfloor N/2\rfloor}(\beta\omega_n+h)\delta_n,$$

where $\{\omega_n\}_{n\in\mathbb{Z}}$ are IID random variables, and then for a local observable f, i.e., a function of τ which depends only on $\tau\cap I$ with I a finite subset of $\mathbb{Z}$, we let

$$\mathbf{E}^{\beta,h}_{\infty,\omega}(f):=\lim_{N\to\infty}\mathbf{E}^{\beta,h}_{N,\omega}(f). \tag{6.1}$$

Existence of the limit, in the localized phase, for almost every disorder realization is proven in [25] (cf. also [9], where a DLR-like point of view is adopted).[4]

The definition of the correlation length ξ contains always some degree of arbitrariness, but conventional wisdom on universality states that the critical properties of ξ, close to a second-order phase transition, are insensitive to the precise definition. There is however a subtlety: in the case of disordered systems there are two possible definitions of correlation lengths, which have no reason to have the same critical behavior. Remaining for definiteness in the framework of our disordered pinning models, one can first of all define a *(disorder-dependent) two-point function* as

$$\mathcal{C}_\omega(k,\ell):=\mathbf{P}^{\beta,h}_{\infty,\omega}(k\in\tau|\ell\in\tau)-\mathbf{P}^{\beta,h}_{\infty,\omega}(k\in\tau). \tag{6.2}$$

In words, $\mathcal{C}_\omega(k,\ell)$ quantifies how much the occurrence of $\ell\in\tau$ influences the occurrence the event $k\in\tau$. It is then natural to define a correlation length ξ as

$$\frac{1}{\xi}:=-\lim_{k\to\infty}\frac{1}{k}\log|\mathcal{C}_\omega(k,0)|, \tag{6.3}$$

provided the limit exists. Note that ξ depends on (β,h) and, in principle, on ω. One can however define a different correlation length, ξ^{av}, as

$$\frac{1}{\xi^{av}}:=-\lim_{k\to\infty}\frac{1}{k}\log\mathbb{E}|\mathcal{C}_\omega(k,0)|. \tag{6.4}$$

[4] One might give a different definition of the infinite-volume Gibbs measure, considering the original system (2.8) defined in $\{1,\dots,N\}$ and taking a the $N\to\infty$ limit of the average of local functions of $\tau\cap I$, with I a finite subset of $\mathbb{N}$. In other words, with the first procedure, Eq. (6.1), we are looking at the system in a window which is situated in the bulk, very far away from both boundaries. On the other hand, the second procedure is relevant if one wants to study the system in the vicinity of one of the two boundaries (and very far away from the other one).

In other words, ξ (respectively, ξ^{av}) is the length over which the two-point function (respectively, the averaged two-point function) decays exponentially. For simplicity, we will call ξ the *typical (or quenched) correlation length*, and ξ^{av} the *average correlation length*, although it is important to keep in mind that ξ^{av} *is not* the disorder-average of ξ (indeed, in Section 6.3 we will see an example where ξ is almost-surely constant but $\xi \neq \xi^{av}$). It is interesting that in the case of the one-dimensional quantum Ising chain with random transverse field studied in [20], the two correlation lengths are believed, on the basis of a renormalization group analysis, to diverge at criticality *with two different critical exponents*.

A simple application of Jensen's inequality shows that $\xi^{av} \geq \xi$. This inequality can be interpreted on the basis of the following intuitive argument. Divide all possible disorder realizations into sets A_m where the empirical average of ω in the region $\{1, \ldots, k\}$ is approximately m. Of course, for $m \neq 0$ A_m is a large deviation-like event of probability $\simeq \exp(-km^2/2)$. Conditionally on A_m, the system sees a defect line which is more attractive (if $m > 0$) or more repulsive (if $m < 0$) than it should and therefore it is more localized (resp. more delocalized) in this region than in the rest of the system. Therefore, conditionally on A_m, we can expect that $\mathcal{C}_\omega(k, 0)$ behaves like $exp(-k/\xi(\beta, h+\beta m))$. In other words, we can argue that (looking only at the exponential behavior)

$$\mathbb{E}\,\mathcal{C}_\omega(k,0) \simeq \int \mathrm{d}m\, e^{-km^2/2} e^{-k/\xi(\beta,h+\beta m)} \simeq e^{k \max_m\{-m^2/2 - 1/\xi(\beta,h+\beta m)\}} \tag{6.5}$$

for k large. Since ξ should diverge when the critical point is approached, it is reasonably decreasing in h so that the value of m which realizes the maximum is strictly negative. On the other hand, when we take the limit without disorder average as in (6.3), the events A_m with $m \neq 0$ cannot contribute, i.e., almost surely they do not occur for k large enough, as follows from the Borel-Cantelli lemma.

6.1 Correlation Length of the Homogeneous Model

In the homogeneous case, $\beta = 0$, the infinite-volume Gibbs measure can be explicitly described (cf. [22, Th. 2.3]): under $\mathbf{P}_\infty^{0,h}(.)$, τ is a homogeneous[5], positively recurrent (for $h > h_c(0) = 0$) renewal on $\mathbb{Z}$ such that

[5] That is, its law is invariant under translation on $\mathbb{Z}$. For instance, $\mathbf{P}_\infty^{0,h}(n, m \in \tau) = \mathbf{P}_\infty^{0,h}(n+k, m+k \in \tau)$ for every $k \in \mathbb{Z}$.

$$\mathbf{P}_{\infty}^{0,h}(\inf\{k>0 : k\in\tau\}=n|0\in\tau)=K(n)e^{-F(0,h)n}e^{h}=:\widetilde{K}_h(n) \tag{6.6}$$

and

$$\mathbf{P}_{\infty}^{0,h}(n\in\tau)=\frac{1}{\sum_{m\in\mathbb{N}}m\widetilde{K}_h(m)}.$$

Note that $\widetilde{K}_h(.)$ is a probability on $\mathbb{N}$ (cf. Eq. (4.1) and the discussion after it) with an exponential tail. What we are interested in is the precise large-n behavior of

$$\mathbf{P}_{\infty}^{0,h}(n\in\tau|0\in\tau)-\frac{1}{\sum_{m\in\mathbb{N}}m\widetilde{K}_h(m)},$$

i.e., a refinement of the renewal theorem (which simply states that this quantity tends to zero for $n\to\infty$).

Let us for a moment widen our scope and consider a homogeneous, positively recurrent renewal, with law $\widetilde{\mathbf{P}}$, such that the law of the distance between two successive points, denoted by $\widetilde{K}(.)$, has exponential tail: say,

$$\lim_{n\to\infty}\frac{1}{n}\log\widetilde{K}(n)=-z<0. \tag{6.7}$$

We *do not* require for the moment that $\widetilde{K}(.)$ is given by (6.6) with $K(.)$ in the class (2.2). It is known (cf. for instance [6, Chapter VII.2] and [38]) that, under condition (6.7), there exist $r>0$ and $C<\infty$ such that

$$\left|\widetilde{\mathbf{P}}(n\in\tau|0\in\tau)-\frac{1}{\sum_{m\in\mathbb{N}}m\widetilde{K}(m)}\right|\le Ce^{-rn}. \tag{6.8}$$

However, the relation between z and the largest possible r in Eq. (6.8), call it r_{max}, is not known in general. A lot of effort has been put by the queuing theory community in investigating this point, and in various special cases it has been proven that $r_{max}\ge z$ (see for instance [7], where power series methods are employed and explicit upper bounds on the prefactor C are given). In even more special cases, for instance when $\widetilde{\mathbf{P}}$ is the law of the return times to a particular state of a Markov chain with some stochastic ordering properties, the optimal result $r_{max}=z$ is proved (for details, see [36, 41], which are based on coupling techniques). However, the equality $r_{max}=z$ cannot be expected in general. In particular, if $\widetilde{K}(.)$ is a geometric distribution,

$$\widetilde{K}(n) = \frac{e^{-nc}}{e^c - 1}$$

with $c > 0$, then one sees easily that the left-hand side of (6.8) vanishes for every $n \in \mathbb{N}$ so that $r_{max} = \infty$, while $z = c$. On the other hand, if for instance $\widetilde{K}(1) = \widetilde{K}(2) = 1/2$ and $\widetilde{K}(n) = 0$ for $n \geq 3$, then $z = +\infty$ while r is finite. These and other nice counter-examples are discussed in [7].

In view of this situation, it is highly non-trivial that, restricting to our original class of renewals, the following holds:

Theorem 6.1. *[23] Let $\widetilde{K}_h(.)$ be given by* (6.6) *with $K(.)$ satisfying* (2.2) *for some $\alpha > 0$ and slowly varying $L(.)$. Then, there exists $h_0 > 0$ such that, for every $0 < h < h_0$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \left| \mathbf{P}_\infty^{0,h}(n \in \tau | 0 \in \tau) - \frac{1}{\sum_{m\in\mathbb{N}} m\widetilde{K}_h(m)} \right| = -F(0,h) \quad (6.9)$$

and, more precisely,

$$\mathbf{P}_\infty^{0,h}(n \in \tau | 0 \in \tau) - \frac{1}{\sum_{m\in\mathbb{N}} m\widetilde{K}_h(m)} \overset{n\to\infty}{\sim} \frac{Q(n)e^{-F(0,h)n}}{4[\sinh(h/2)]^2} \quad (6.10)$$

with $Q(.)$ such that $\sum_{j=1}^n Q(j) \overset{n\to\infty}{\sim} L(n)/(\alpha n^\alpha)$.

It is important to emphasize that, even under assumption (6.6), this result would be false without the restriction of h small.

In the light of (6.9), it is quite natural to expect (and in some case this can be proven, see Section 6.3) that in presence of disorder ξ is still proportional to the inverse of the free energy, at least close to the critical point. But then, what about ξ^{av}?

6.2 μ versus F

To answer this question, we abandon for a while the correlation length and we discuss the relation between free energy and another quantity which, due to lack of a standard name, we will call simply μ. This was first introduced, to my knowledge, in [3] in the context of random heteropolymers:

$$\mu(\beta, h) = -\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\frac{1}{Z_{N,\omega}(\beta, h)} \right] \quad (6.11)$$

Existence of the limit in our context is easily proven by super-additivity of $\log Z_{N,\omega}(\beta,h)$ (see [25, Th. 2.5]). An argument similar to (2.12) gives immediately $\mu \geq 0$ while a simple application of Jensen's inequality shows that $\mu(\beta,h) \leq F(\beta,h)$. However, much more than this is true:

Theorem 6.2. *[41] For every $\beta > 0$ there exists $0 < c_3(\beta), c_4(\beta) < \infty$ such that*

$$0 < c_3(\beta)\frac{F(\beta,h)^2}{\partial_h F(\beta,h)} < \mu(\beta,h) < F(\beta,h) \tag{6.12}$$

if $0 < h - h_c(\beta) \leq c_4(\beta)$.

In particular, the bounds in (6.12) show that also μ vanishes continuously at the critical point, like the free energy. If we call η_F and η_μ the critical exponents associated to the vanishing of F and μ for $h \to h_c(\beta)^+$, Theorem 6.2 implies the following bounds:

$$(2 \leq)\eta_F \leq \eta_\mu \leq \eta_F + 1, \tag{6.13}$$

the inequality in parentheses being valid for $\beta > 0$ thanks to Theorem 5.8. Just to give a flavor of why μ is relevant in the description of the system let us cite the following result. Define first of all Δ_N as the largest gap between points of τ in the system of length N:

$$\Delta_N := \max_{1 \leq i < j \leq N} \{|i-j| : i \in \tau, j \in \tau, \{i+1, \ldots, j-1\} \cap \tau = \emptyset\}. \tag{6.14}$$

Then,

Theorem 6.3. *[25] Let $(\beta,h) \in \mathcal{L}$. For every $\epsilon > 0$,*

$$\lim_{N\to\infty} \mathbf{P}^{\beta,h}_{N,\omega}\left(\frac{1-\epsilon}{\mu(\beta,h)} \leq \frac{\Delta_N}{\log N} \leq \frac{1+\epsilon}{\mu(\beta,h)}\right) = 1 \quad \textit{in probability.} \tag{6.15}$$

6.3 Correlation Lengths and Free Energy

To my knowledge, the only case where ξ and ξ^{av} can be fully characterized even in presence of disorder is the one where $K(.)$ is the law of the first return to zero of the one-dimensional SRW conditioned to be non-negative. In other words, let $\{S_n\}_{n=0,1,\ldots}$ be the SRW on $\mathbb{Z}$ started at $S_0 = 0$ and let $\mathbf{P}^{SRW}(.)$ denote its law. We define $K^{SRW,+}(n) := \mathbf{P}^{SRW}(\inf\{k > 0 : S_k = 0\} = 2n | S_i \geq 0 \ \forall i)$. Go back to Section 2.6 for a motivation of this example as a model of wetting of a $(1+1)$-dimensional substrate. In this case, one has the following

Theorem 6.4. *[41] Let $K(.) = K^{SRW,+}(.)$ and $\ell \in \mathbb{Z}$. For every $\beta \geq 0$ and $h > h_c(\beta)$,*

$$\frac{1}{\xi^{av}} = -\lim_{k\to\infty} \frac{1}{k} \log \mathbb{E}\, \mathcal{C}_\omega(\ell + k, k) = \mu(\beta, h) \tag{6.16}$$

and, $\mathbb{P}(\mathrm{d}\omega)$–a.s.,

$$\frac{1}{\xi} = -\lim_{k\to\infty} \frac{1}{k} \log \mathcal{C}_\omega(\ell + k, k) = F(\beta, h). \tag{6.17}$$

With respect to Theorem 6.1, this result is much less sharp in that it catches only the exponential behavior of the two-point function. However, note that in Theorem 6.4 $h - h_c(\beta)$ is not required to be small as in Theorem 6.1. Note also that in Eqs. (6.16), (6.17) we have not taken the absolute value of $C_\omega(\ell + k, k)$: this is because, in this particular case, one can prove that this quantity is non-negative [41]. Finally observe that, in view of (6.12), the two correlation lengths are different. It would be extremely interesting to know whether the two associated critical exponents η_F, η_μ coincide or not.

Remark 6.5. Theorem 6.4 does not coincide exactly with [41, Th. 3.5], e.g., because in the latter $\mathbf{P}^{\beta,h}_{\infty,\omega}(.)$ is the infinite-volume Gibbs measure obtained from the system defined in $\{1, \ldots, N\}$ letting $N \to \infty$ (cf. footnote 4). However, the proof of [41] extends without difficulties to the result we stated above. We remark also that the theorem holds as well in the case where $K(n) = K^{SRW}(n) := \mathbf{P}^{SRW}(\inf\{k > 0 : S_k = 0\} = 2n)$, i.e., the law of the first return to zero of the unconditioned SRW. This follows from the discussion in Section 2.6 and from the fact that $K^{SRW}(n) = 2K^{SRW,+}(n)$.

Proof of Theorem 6.4 (sketch). The proof of Theorem 6.4 is based on a coupling argument. For simplicity let $\mathbf{P}^+(.) := \mathbf{P}^{SRW}(.|S \geq 0)$. One can then rewrite the two-point function (6.2) as

$$\mathcal{C}_\omega(k, \ell) = \lim_{N\to\infty} \frac{1}{Z_{N,\omega}(\beta, h)^2} \times$$
$$\times \mathbf{E}^{+,\otimes 2}\left[e^{\sum_{n=-N/2}^{N/2} (\beta\omega_n + h)(\mathbf{1}_{\{S_n^{(1)}=0\}} + \mathbf{1}_{\{S_n^{(2)}=0\}})} \left(\mathbf{1}_{\{S_k^{(1)}=0\}} - \mathbf{1}_{\{S_k^{(2)}=0\}} \right) | S_\ell^{(1)} = 0 \right], \tag{6.18}$$

where $S^{(1)}, S^{(2)}$ are independent with law $\mathbf{P}^+$. Since the SRW conditioned to be non-negative is a Markov chain, the expectation in the right-hand side clearly vanishes if we condition on the event that there

exists $\ell < i < k$ such that $S_i^{(1)} = S_i^{(2)}$. But (and here we use explicitly the condition $S_i \geq 0$ and that two SRW trajectories which cross each other do necessarily intersect), if the complementary event happens then either $S^{(1)}$ or $S^{(2)}$ has no zeros in the interval $\{\ell+1, \dots, k-1\}$. As a consequence, one obtains

$$\mathbb{E}\,\mathcal{C}_\omega(k,0) \leq 2\mathbb{E}\,\mathbf{P}_{\infty,\omega}^{\beta,h}(\tau \cap \{1,\dots,k-1\} = \emptyset) \tag{6.19}$$

and it is not difficult to deduce from (6.11) that this probability vanishes like $\exp(-k\mu(\beta,h))$ for $k \to \infty$. For the opposite bound and for the proof of (6.17) we refer to [41].

In the general case where $\mathbf{P}$ is not necessarily the law of the returns of the SRW (or, in general, of any Markov chain), the available results on correlation lengths in presence of disorder are much less sharp and, above all, only correlation length *upper bounds* are known. At present, the best one can prove in general about average correlation length is the following:

Theorem 6.6. *[42] Let $\epsilon > 0$ and $(\beta,h) \in \mathcal{L}$. There exists $C_1 := C_1(\epsilon,\beta,h) > 0$ such that, for every $k \in \mathbb{N}$,*

$$\mathbb{E}\,|\mathcal{C}_\omega(\ell+k,\ell)| \leq \frac{1}{C_1\mu(\beta,h)^{1/C_1}} \exp\left(-k\,C_1\,\mu(\beta,h)^{1+\epsilon}\right). \tag{6.20}$$

The constant $C_1(\epsilon,\beta,h)$ does not vanish at the critical line: for every bounded subset $B \subset \mathcal{L}$ one has $\inf_{(\beta,h)\in B} C_1(\epsilon,\beta,h) \geq C_1(B,\epsilon) > 0$.

Remark 6.7. The necessity of introducing $\epsilon > 0$ (i.e., of weakening the upper bound with respect to the expected one) is probably of technical nature, as appears from the fact that for $\beta = 0$ Theorem 6.6 does not reproduce the sharp results (6.9) which hold for the homogeneous case.

Observe that Theorem 6.6 is more than just an upper bound on ξ^{av}. Indeed, thanks to the bound on the prefactor in front of the exponential, Eq. (6.20) says that the exponential decay, with rate at least of order $\mu^{1+\epsilon}$, starts as soon as $k \gg \mu^{-1-\epsilon}|\log\mu|$. This observation reinforces the meaning of Eq. (6.20) as an upper bound of order μ^{-1} on the correlation length of disorder-averaged correlations functions.

About the typical correlation length the following can be proven:

Theorem 6.8. *[42] Let $\epsilon > 0$ and $(\beta,h) \in \mathcal{L}$. One has for every $k \in \mathbb{N}$*

$$|\mathcal{C}_\omega(k,0)| \leq C_2(\omega)\exp\left(-k\,C_1\,F(\beta,h)^{1+\epsilon}\right), \tag{6.21}$$

where C_1 is as in Theorem 6.6, while $C_2(\omega) := C_2(\omega,\epsilon,\beta,h)$ is an almost surely finite random variable.

The proof of Theorems 6.6 and 6.8 relies on a rather involved coupling/comparison argument. In simple (and imprecise) words, one first approximates $K(.)$ with a new law $\widetilde{K}(.)$ which is the law of the returns to zero of a Markov process with continuous trajectories (defined in terms of a Bessel process), and at that point the coupling argument of last section can be applied. We refer to [42] for full details.

A Two Tauberian Results

For completeness, we include without proof two Tauberian theorems (i.e., results about the relation between the asymptotic behavior of a function and of its Laplace transform) which we used in Section 5.5. Given a function $Q : \mathbb{N} \to \mathbb{R}$, we define for $s \in \mathbb{R}$

$$\hat{Q}(s) := \sum_{n \in \mathbb{N}} e^{-ns} Q(n)$$

whenever the sum converges.

We begin with a (quite intuitive) fact:

Theorem A.1. *[8, Proposition 1.5.8] If $\ell(.)$ is slowly varying and $\gamma > -1$ then*

$$\sum_{n=1}^{N} n^{\gamma} \ell(n) \overset{N \to \infty}{\sim} \frac{N^{\gamma+1}}{\gamma+1} \ell(N). \tag{A.1}$$

Next we state Karamata's Tauberian theorem [8, Th. 1.7.1] which for our purposes may be formulated as follows:

Theorem A.2. *Assume that $Q(n) \geq 0$ for every $n \in \mathbb{N}$, that $\ell(.)$ is slowly varying and that $\rho \geq 0$. The following are equivalent:*

$$\hat{Q}(s) \overset{s \searrow 0}{\sim} \frac{\ell(1/s)}{s^{\rho}} \tag{A.2}$$

and

$$\sum_{n=1}^{N} Q(n) \overset{N \to \infty}{\sim} N^{\rho} \frac{\ell(N)}{\Gamma(1+\rho)}. \tag{A.3}$$

Recall that the function $\Gamma(z)$ can be defined, for $z > 0$, as

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} \mathrm{d}t.$$

Finally, a theorem relating the Laplace transform of a law on the half-line to its integrated tail (cf. [8, Corollary 8.1.7]):

Theorem A.3. *Let X be an integer-valued random variables with law* $\mathbf{P}$ *and* $Q(n) := \mathbf{P}(X = n)$, $\ell(.)$ *a slowly varying function and* $0 \leq \alpha < 1$. *The following are equivalent:*

$$1 - \hat{Q}(s) \overset{s\searrow 0}{\sim} s^{\alpha}\ell(1/s) \tag{A.4}$$

and

$$\mathbf{P}(X > n) = \sum_{j>n} Q(j) \overset{n\to\infty}{\sim} \frac{\ell(n)}{n^{\alpha}\Gamma(1-\alpha)}. \tag{A.5}$$

Acknowledgments

I would like to thank Roman Kotecký for organizing the Prague Summer School on Mathematical Statistical Mechanics and for inviting me to give a course. Learning and teaching there was an extremely stimulating experience.

Some of the results described in these notes are based on joint work with Giambattista Giacomin, to whom I am grateful for introducing me to this subject, for countless motivating conversations, and also for communicating to me the results of [23] prior to publication.

This work was supported in part by the GIP-ANR project JC05_42461 (*POLINTBIO*) and my presence at the school was made possible thanks to the support from the ESF-program "Phase Transitions and Fluctuation Phenomena for Random Dynamics in Spatially Extended Systems".

References

1. M. Aizenman, R. Sims, S. L. Starr, *Extended variational principle for the Sherrington-Kirkpatrick spin-glass model*, Phys. Rev. B **68**, 214403 (2003).
2. M. Aizenman and J. Wehr, *Rounding effects of quenched randomness on first–order phase transitions*, Comm. Math. Phys. **130** (1990), 489–528.
3. S. Albeverio and X. Y. Zhou, *Free energy and some sample path properties of a random walk with random potential*, J. Statist. Phys. **83** (1996), 573–622.
4. K. S. Alexander, *The effect of disorder on polymer depinning transitions*, Commun. Math. Phys. **279**, 117–146 (2008).

5. K. S. Alexander, V. Sidoravicius, *Pinning of polymers and interfaces by random potentials*, Ann. Appl. Probab. **16**, 636–669 (2006).
6. S. Asmussen, *Applied Probability and Queues*, 2nd ed., Springer-Verlag, New York, 2003.
7. K. S. Berenhaut, R. B. Lund, *Renewal convergence rates for DHR and NWU lifetimes*, Probab. Engrg. Inform. Sci. **16**, 67–84 (2002).
8. N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular Variation*, Cambridge University Press, Cambridge (1987).
9. M. Biskup, F. den Hollander, *A heteropolymer near a linear interface*, Ann. Appl. Probab. **9**, 668–687 (1999).
10. T. Bodineau, G. Giacomin, *On the localization transition of random copolymers near selective interfaces*, J. Stat. Phys. **117**, 801–818 (2004).
11. E. Bolthausen, F. den Hollander, *Localization transition for a polymer near an interface*, Ann. Probab. **25**, 1334–1366 (1997).
12. F. Caravenna, G. Giacomin and M. Gubinelli *A numerical approach to copolymers at selective interfaces*, J. Stat. Phys. **122**, 799–832 (2006).
13. J. T. Chayes, L. Chayes, D. S. Fisher and T. Spencer, *Correlation Length Bounds for Disordered Ising Ferromagnets*, Commun. Math. Phys. **120**, 501–523 (1989).
14. J. T. Chayes, L. Chayes, D. S. Fisher and T. Spencer, *Finite-size scaling and correlation lengths for disordered systems*, Phys. Rev. Lett. **57**, 2999 (1986).
15. D. Cule, T. Hwa, *Denaturation of Heterogeneous DNA*, Phys. Rev. Lett. **79**, 2375–2378 (1997).
16. B. Derrida, G. Giacomin, H. Lacoin, F. L. Toninelli, *Fractional moment bounds and disorder relevance for pinning models*, Commun. Math. Phys., to appear arxiv.org/abs/0712.2515 [math.PR].
17. B. Derrida, V. Hakim and J. Vannimenius, *Effect of disorder on two-dimensional wetting*, J. Statist. Phys. **66** (1992), 1189–1213.
18. R. A. Doney, *One-sided large deviation and renewal theorems in the case of infinite mean*, Probab. Theory Rel. Fields **107**, 451–465 (1997).
19. W. Feller, *An introduction to probability theory and its applications*, vol. 1, 2^{nd} ed., John Wiley & Sons Inc. (1966).
20. D. S. Fisher, *Random transverse field Ising spin chains*, Phys. Rev. Lett. **69** (1992), 534–537.
21. G. Forgacs, J. M. Luck, Th. M. Nieuwenhuizen and H. Orland, *Wetting of a Disordered Substrate: Exact Critical behavior in Two Dimensions*, Phys. Rev. Lett. **57** (1986), 2184–2187.
22. G. Giacomin, *Random polymer models*, Imperial College Press, Imperial College Press, World Scientific (2007).
23. G. Giacomin, *Renewal convergence rates and correlation decay for homogeneous pinning models*, Elect. J. Probab. **13**, 513–529 (2008).
24. G. Giacomin, F. L. Toninelli, *Estimates on path delocalization for copolymers at selective interfaces*, Probab. Theory Rel. Fields **133**, 464–482 (2005).

25. G. Giacomin and F. L. Toninelli, *The localized phase of disordered copolymers with adsorption*, ALEA **1** (2006), 149–180.
26. G. Giacomin and F. L. Toninelli, *Smoothing effect of quenched disorder on polymer depinning transitions*, Commun. Math. Phys. **266**, 1–16 (2006).
27. G. Giacomin, F. L. Toninelli, *Smoothing of Depinning Transitions for Directed Polymers with Quenched Disorder*, Phys. Rev. Lett **96**, 060702 (2006).
28. G. Giacomin, F. L. Toninelli, *On the irrelevant disorder regime of pinning models*, Ann. Probab., to appear arXiv:0707.3340v1 [math.PR]
29. F. Guerra, *Replica Broken Bounds in the Mean Field Spin Glass Model*, Commun. Math. Phys. **233**, 1–12 (2003).
30. F. Guerra, *Sum rules for the free energy in the mean field spin glass model*, in *Mathematical Physics in Mathematics and Physics: Quantum and Operator Algebraic Aspects*, Fields Inst. Commun. **30**, AMS, 2001.
31. F. Guerra, F. L. Toninelli, *Quadratic replica coupling for the Sherrington-Kirkpatrick mean field spin glass model*, J. Math. Phys. **43**, 3704–3716 (2002).
32. F. Guerra, F. L. Toninelli, *The Thermodynamic Limit in Mean Field Spin Glass Models*, Commun. Math. Phys. **230**, 71–79 (2002).
33. A. B. Harris, *Effect of Random Defects on the Critical Behaviour of Ising Models*, J. Phys. C **7**, 1671–1692 (1974).
34. N. C. Jain, W. E. Pruitt, *The Range of Rando Walk*, in *Proceedings of the Sixth Berkeley Simposium on Mathematical Statistics and Probability* (Univ. California, Berkeley, Calif., 1970/71), Vol. III: Probability Theory, pp. 31–50, Univ. California Press, Berkeley, Cakuf., 1972.
35. Y. Kafri, D. Mukamel and L. Peliti *Why is the DNA denaturation transition first order?*, Phys. Rev. Lett. **85**, 4988–4991 (2000).
36. R. B. Lund, R. L. Tweedie, *Geometric convergence rates for stochastically ordered Markov chains*, Math. Oper. Res. **21**, 182–194 (1996).
37. C. Monthus, *On the localization of random heteropolymers at the interface between two selective solvents*, Eur. Phys. J. B **13**, 111–130 (2000).
38. P. Ney, *A refinement of the coupling method in renewal theory*, Stochastic Process. Appl. **11**, 11–26 (1981).
39. M. Talagrand, *The Parisi Formula*, Ann. Math. **163**, 221–263 (2006).
40. M. Talagrand, *Spin glasses, a Challenge for Mathematicians*, Springer-Verlag (2003).
41. F. L. Toninelli, *Critical properties and finite-size estimates for the depinning transition of directed random polymers*, J. Stat. Phys. **126**, 1025–1044 (2007).
42. F. L. Toninelli, *Correlation lengths for random polymer models and for some renewal sequences*, Electron. J. Probab. **12**, 613–636 (2007).
43. F. L. Toninelli, *A replica-coupling approach to disordered pinning models*, Commun. Math. Phys. **280**, 389–401 (2008).
44. F. L. Toninelli, *Disordered pinning models and copolymers: beyond annealed bounds*, Ann. Appl. Probab. **18**, 1569–1587 (2008).

Metastability

Anton Bovier

Institut für Angewandte Mathematik
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstrasse 6, 53115 Bonn, Germany
anton.bovier@uni-bonn.de

Contents

R. Kotecký (ed.), *Methods of Contemporary Mathematical Statistical Physics*,
Lecture Notes in Mathematics 1970, DOI: 10.1007/978-3-540-92796-9_4,

1 Introduction

In these lectures we will discuss Markov processes with a particular interest for a phenomenon called *metastability*. Basically this refers to the existence of two or more time-scales over which the system shows very different behaviour: on the short time scale, the systems reaches quickly a "pseudo-equilibrium" and remains effectively in a restricted subset of the available phase space; the particular pseudo-equilibrium that is reached will depend on the initial conditions. However, when observed on the longer time scale, one will occasionally observe *transitions* from one such pseudo-equilibrium to another one. In many cases (as we will see) there exists one particular time scale for each such pseudo-equilibrium; in other cases of interest, several, or even many, such distinct pseudo-equilibria exist having the same time scale of exit. Mathematically speaking, our interest is to derive the (statistical) properties of the process on these long time scales from the given description of the process on the microscopic time scale. In principle, our aim should be an effective model for the motion at the long time scale on a coarse grained state space; in fact, disregarding fast motion leads us naturally to consider a reduced state space that may be labeled in some way by the quasi-equilibria.

The type of situation we sketched above occurs in many situations in nature. The classical example is of course the phenomenon of metastability in *phase transitions*: if a (sufficiently pure) container of water is cooled below freezing temperature, it may remain in the liquid state for a rather long period of time, but at some moment the entire container freezes extremely rapidly. In reality, this moment is of course mostly triggered by some slight external perturbation. Another example of the same phenomenon occurs in the dynamics of large bio-molecules, such as proteins. Such molecules frequently have several possible spatial *conformations*, transitions between which occur sporadically on often very long time scales. Another classical example is metastability in chemical reactions. Here reactants oscillate between several possible chemical compositions, sometimes nicely distinguished by different colours. This example was instrumental in the development of stochastic models for metastability by Eyring, Kramers and others [21, 30]. Today, metastable effects are invoked to explain a variety of diverse phenomena such as changes in global climate systems both on earth (ice-ages) and on Mars (liquid water presence), structural transitions on eco- and oeco systems, to name just a few examples.

Most modeling approaches attribute metastability to the presence of some sort of randomness in the underlying dynamics. Indeed, in

the context of purely deterministic systems, once several equilibrium positions for the dynamics exist, transitions between such equilibria are impossible. It is then thought that metastable effects occur due to the presence of (small) random perturbations that should reflect the influence of unresolved degrees of freedom on very fast scales.

Mathematically, metastability is studied in a number of contexts of which we mention the following:

(i) **Small random perturbations of dynamical systems.** Here one considers a classical dynamical system in $\mathbb{R}^d$ with some added small stochastic noise term. This leads to a stochastic differential equation of the type

$$dx_\epsilon(t) = f_\epsilon(x_\epsilon(t))dt + \sqrt{\epsilon} g_\epsilon(x_\epsilon(t))dW(t) \tag{1.1}$$

Such systems have been extensively investigated e.g. in the work of Freidlin and Wentzell [23] and Kifer [28]. They have their origin in the work of Kramers [30].

(ii) **Markov chains with exponentially small transition rates.** Here we are dealing with Markov chains with discrete state space that are almost deterministic in the sense that the transition probabilities are either exponentially close to one or exponentially close to zero, in some small parameter ϵ. Such systems emerge in the analysis of Wentzell and Freidlin and are studied there. They found renewed interest in the context of low temperature dynamics for lattice models in statistical mechanics [35, 36, 1] and also in the analysis of stochastic algorithms for the solution of optimisation problems ("simulated annealing") [14, 13]. Recent result using the methods outlined here can be found in [11, 6].

(iii) **Glauber dynamics of mean field [12, 32, 22, 7] or lattice [37] spin systems.** Metastability in stochastic dynamics of spin systems is not restricted to the zero temperature limit, but happens whenever there is a first order phases transition. At finite temperature, this is much harder to analyse in general. The reason is that it is no longer true that the process on the micro-scale is close to deterministic, but that such a statement may at best be meaningful on a coarse grained scale. Mean field models lend themselves to such a coarse graining in a particularly nice way, and in many cases it is possible to construct an effective coarse grained Markovian dynamics that then is in some sense similar to the problems mentioned in (i).

The traditional methods to analyse such systems are

(a) **Large deviations.** Wentzell and Freidlin introduced the method of large deviations on path space in order to obtain a rigorous analysis of the probability for the deviations of solutions of the stochastic differential equations (1.1) from the solutions of the deterministic limiting equations. This method has proven very robust and has been adapted to all of the other contexts. The price to pay for generality is limited precision. In general, only the exponential rates of probabilities can be computed precisely. Frequently this is good enough in applications, but sometimes more precise results are desirable. In certain cases, refined estimates could, however, be obtained [19].

(b) **Asymptotic perturbation theory.** As we will see in detail in the course of these lectures, many key quantities of interest concerning Markov processes can be characterized as solutions of certain systems of linear equations, that are, or are structurally similar to, boundary value problems in partial differential equations. In particular cases of stochastic differential equations with small noise, or discrete versions thereof, one may use methods from perturbation theory of linear differential operators with the variance of the noise playing the rôle of a small parameter. This has been used widely in the physics literature on the subject (see e.g. the book by Kolokoltsov [29] for detailed discussions and further reference), however, due to certain analytic difficulties, with the exception of some very particular cases, a rigorous justification of these methods was not given. A further shortcoming of the method is that it depends heavily on the particular types of Markov processes studied and does not seem to be universally applicable. Very recently, Helffer, Nier and Klein have been able to develop a new analytic approach that allows to develop rigorous asymptotic expansion for the small eigenvalues for diffusion processes [26, 25, 34].

(c) **Spectral and variational methods.** Very early on it was noted that there should be a clear signature of metastability in the nature of the generator (or transition matrix) of the Markov process considered. To see this, note that if the Markov process was effectively reducible, i.e. had instead of quasi invariant sets there were truly invariant sets, then the generator would have a degenerate eigenvalue zero with multiplicity equal to the number of invariant sets. Moreover, the eigenfunctions could be chosen as the indicator functions of these sets. It is natural to believe that a perturbed version of this picture remains true in the metastable setting. The computation

of small eigenvalues and "spectral gaps" has thus been a frequent theme in the subject. Computations of eigenvalues can be done using variational representations of eigenvalues, and a number of rather precise results could be achieved in this way, e.g. in the work of Mathieu [31] and Miclo [33].

In these lectures I will explain an approach to metastability that is in some sense mixing ideas from (ii) and (iii) and that proves to be applicable in a wide variety of situations. One of its goals is to obtain a precise characterization of metastability in terms of spectral characteristics, and in particular a quantitatively precise relation between eigenvalues and physical quantities such as exit times from metastable domains. The main novel idea in this approach, that was developed in collaboration with M. Eckhoff, V. Gayrard, and M. Klein over the last years [7, 8, 9, 10] (see also the reviews [3, 4]) is the systematic use of the so called "Newtonian capacity", a fundamental object in potential theory, and its variational representation. This will allow us to get in a rigorous way results that are almost as precise as those obtained from perturbation theory in a rather general context. In particular, we will see that certain structural relations between capacities, exit times and spectral characteristics hold without further model assumptions under some reasonable assumptions on what is to be understood by the notion of metastability.

In these lectures I will focus on the general methodology of this approach. In Sections 3 and 4 I outline the universal relations between capacity and metastable exit times in the context of discrete Markov chains (where this approach is fully developed), and in the same context the relation to spectral theory is explained in Section 6. These results are "model independent", in certain sense. To apply these to specific models, one needs to compute certain capacities. Here, too, we have developed a rather clear strategy of how to do this, which is explained in Section 6 and exemplified in the case of the Curie-Weiss model.

The real test of any method comes when it is applied in non-trivial examples. The lecture notes of Frank den Hollander in this volume [20] present some of these in the context of Glauber and Kawasaki dynamics of lattice gases. Besides apparent successes, there remain many challenges and some directions of ongoing research will be exposed in his lectures.

Acknowledgment. The results described in these lectures outline the approach to metastability developed in collaboration with M. Eckhoff, V. Gayrard, and M. Klein, and further elaborated with A. Bianchi, D. Ioffe, F. den Hollander, F. Manzo, F. Nardi, and C. Spitoni.

Section 5.4 contains unpublished material based on a collaboration with A. Bianchi and D. Ioffe. This research is supported in part by the German Research Council in the Dutch-German Bilateral Research Group "Random Spatial Models from Physics and Biology".

2 Basic Notions from the Theory of Markov Processes

A stochastic process $\{X_t\}_{t\in I}$, $X_t \in \Gamma$ is called a Markov process with *index set* I and *state space* Γ, if, for any collection $t_1 < \cdots < t_n < t \in I$,

$$\mathbb{P}\left[X_t \in \mathcal{A} | X_{t_n} = x_n, \ldots, X_{t_1} = x_1\right] = \mathbb{P}\left[X_t \in \mathcal{A} | X_{t_n} = x_n\right] \tag{2.1}$$

for any Borel set $\mathcal{A} \in \mathcal{B}(\Gamma)$. Here I is always an ordered set, in fact either $\mathbb{N}$ or $\mathbb{R}$. In the former case we call call the process a discrete time Markov chain, the second case is referred to as a continuous time Markov process. A further distinction concerns the nature of the state space Γ. This may be finite, countable, or uncountable ('continuous').

A key quantity in all cases is the family of probability measures, $p(s,t,x,\cdot)$, on $(\Gamma, \mathcal{B}(\Gamma))$,

$$p(s,t,x,\mathcal{A}) \equiv \mathbb{P}\left(X_t \in \mathcal{A} | X_s = x\right), \tag{2.2}$$

for any Borel set $\mathcal{A} \in \mathcal{B}(\Gamma)$. By (2.1), $p(t,s,x,\cdot)$ determines uniquely the law of the Markov process. In fact, any family of probability measures $p(s,t,x,\cdot)$ satisfying

$$p(s,s,x,\cdot) = \delta_x(\cdot) \tag{2.3}$$

and the relation for $s < t' < t$,

$$p(s,t,x,\cdot) = \int p(s,t',x,dz) p(t',t,z,\cdot) \tag{2.4}$$

defines a Markov process. If $p(s,t,x,\cdot)$ is a function of $t-s$ only, we call the Markov process *time-homogeneous* and set

$$p(s,t,x,\cdot) \equiv p_{t-s}(x,\cdot) \tag{2.5}$$

We will only be concerned with time-homogeneous Markov processes henceforth. In the case of discrete time the transition kernel is fully determined by the one-step transition probabilities, called transition matrix in the discrete space case,

$$p(x,\cdot) \equiv p_1(x,\cdot) \tag{2.6}$$

If space is discrete, we can of course simply specify the atoms, $p(x, y)$, of this measure; this object is then called the *transition matrix*.

Property (2.4) is often called the *semi-group* property and the transition kernel $p_t(x, \cdot)$ is called a *Markov semi-group*. In continuous time, one defines the *generator* (of the semi-group)[1]

$$L \equiv \lim_{t \downarrow 0} t^{-1}(1 - p_t) \tag{2.7}$$

It then follows that conversely

$$p_t = e^{-tL} \tag{2.8}$$

We will find it sometimes convenient to define a "generator" also in the discrete time case by setting

$$L \equiv 1 - p_1 \tag{2.9}$$

We will frequently think of p_t and L as operators acting on functions f on Γ as

$$p_t f(x) \equiv \int_\Gamma p_t(x, dy) f(y) \tag{2.10}$$

respectively on measures ρ on Γ, via

$$\rho p_t(\cdot) \equiv \int_\Gamma \rho(dx) p_t(x, \cdot) \tag{2.11}$$

If $\rho_0(\cdot) = \mathbb{P}(X_0 \in \cdot)$, then

$$\rho_0 p_t(\cdot) \equiv \rho_t(\cdot) = \mathbb{P}(X_t \in \cdot) \tag{2.12}$$

ρ_t is called the law of the process at time t started in ρ at time 0. It is easy to see from the semi-group property that ρ_t satisfies the equation

$$\frac{\partial}{\partial t} \rho_t(x, \cdot) = -\rho_t L(x, \cdot) \tag{2.13}$$

resp., in the discrete time case

$$\rho_{t+1}(x, \cdot) = -\rho_t L(x, \cdot) \tag{2.14}$$

This equation is called the *Focker-Planck* equation. A probability measure μ on Γ is called an *invariant measure* for the Markov process X_t if it is a stationary solution of (2.13), i.e. if

$$\mu p_t = \mu \tag{2.15}$$

[1] In the literature, one usually defines the generator with an extra minus sign. I prefer to work with positive operators.

for all $t \in I$. Note that (2.15) is equivalent to demanding that

$$\mu L = 0 \tag{2.16}$$

A priori the natural function space for the action of our operators is $L_\infty(\Gamma)$ for the action from the left, and locally finite measures for the action on the right. Given an invariant measure μ, there is, however, also a natural extension to the space $L_2(\Gamma, \mu)$. In fact, p_t is a *contraction* on this space, and L is a positive operator. To see this, just use the Schwartz inequality to show that

$$\int \mu(dx) \left(\int p_t(x, dy) f(y) \right)^2 \leq \mu(dx) \int p_t(x, dy) f(y)^2 = \int \mu(dy) f(y)^2 \tag{2.17}$$

L is in general not a bounded operator in L_2, and its domain is sometimes just a dense subspaces.

Within this L_2-theory, it is natural to define the adjoint operators p_t^* and L^* via

$$\int \mu(dx) g(x) p_t^* f(x) \equiv \int \mu(dx) f(x) p_t^* g(x) \tag{2.18}$$

respectively

$$\int \mu(dx) g(x) L^* f(x) \equiv \int \mu(dx) f(x) L g(x) \tag{2.19}$$

for any pair of functions $f, g \in L_2(\Gamma, \mu)$. We leave it as an exercise to show that p_t^* and L^* are Markov semi-groups, resp. generators, whenever μ is an invariant measure. Thus they define an *adjoint* or reverse process. In the course of these lectures we will mainly be concerned with the situation where p_t and L are *self-adjoint*, i.e. when $p_t = p_t^*$ and $L = L^*$. This will entrain a number of substantial simplifications. Results on the general case can often be obtained by comparison with symmetrized processes, e.g. the process generated by $(L + L^*)/2$. Note that whenever a Markov generator is self-adjoint with respect to a measure μ, then this measure is invariant (Exercise!). We call Markov processes whose generator is self-adjoint with respect to some probability measure *reversible*. The invariant measure is then often called the *reversible measure* (although I find this expression abusive; symmetrizing measure would be more appropriate).

Working with reversible Markov chains brings the advantage to make full use of the theory of self-adjoint operators, which gives far richer results then in the general case. In many applications one can

work by choice with reversible Markov processes, so that in practical terms this restriction is not too dramatic.

Hitting Times

Henceforth we denote by $\mathbb{P}_x$ the law of the process conditioned on $X_0 = x$. For any (measurable) set $D \subset \Gamma$ we define the hitting time τ_D as

$$\tau_D \equiv \inf (t > 0 : X_t \in D) \tag{2.20}$$

Note that τ_D is a *stopping time*, i.e. the random variable τ_D depends only on the behaviour of X_t for $t \leq \tau_D$. Denoting by $\mathcal{F}_t$ sigma-algebra generated by $\{X_s\}_{0 \leq s \leq t}$, we may say that the event $\{\tau_D \leq t\}$ is measurable with respect to $\mathcal{F}_t$.

3 Discrete Space, Discrete Time Markov Chains

We will now turn to our main tools for the analysis of metastable systems. To avoid technical complications and to focus on the key ideas, we will first consider only the case of discrete (or even finite) state space and discrete time (the latter is no restriction). We set $p_1(x, y) = p(x, y)$. We will also assume that our Markov chain is irreducible, i.e. that for any $x, y \in \Gamma$, there is $t \in \mathbb{N}$ such that $p_t(x, y) > 0$. If in addition Γ is finite, this implies the existence of a unique invariant (probability) measure μ. We will also assume the our Markov chain is reversible.

3.1 Equilibrium Potential, Equilibrium Measure, and Capacity

Given two disjoint subsets A, D, of Γ, and $x \in \Gamma$, we are interested in

$$\mathbb{P}_x[\tau_A < \tau_D] \tag{3.1}$$

One of our first, and as we will see main tasks is to compute such probabilities. We consider first the case of discrete time and space.

If $x \notin A \cup D$, we make the elementary observation that the first step away leads either to D, and the event $\{\tau_A < \tau_D\}$ fails to happen, or to A, in which case the event happens, or to another point $y \notin A \cup D$, in which case the event happens with probability $\mathbb{P}_y[\tau_A < \tau_D]$. Thus

$$\mathbb{P}_x[\tau_A < \tau_D] = \sum_{y \in A} p(x, y) + \sum_{y \notin A \cup D} p(x, y) \mathbb{P}_y[\tau_A < \tau_D] \tag{3.2}$$

We call an equation based on this reasoning a *forward* equation. Note that we can write this in a nicer form if we introduce the function

$$h_{A,D}(x) = \begin{cases} \mathbb{P}_x[\tau_A < \tau_D], & \text{if } x \notin A \cup D \\ 1, & \text{if } x \in A \\ 0, & \text{if } x \in D \end{cases} \tag{3.3}$$

Then (3.2) implies that for $x \notin A \cup D$,

$$h_{A,D}(x) = \sum_{y \in \Gamma} p(x,y) h_{A,D}(y) \tag{3.4}$$

In other words, the function $h_{A,D}$ solves the *boundary value problem*

$$\begin{aligned} Lh_{A,D}(x) &= 0, \quad x \in \Gamma \backslash (A \cup D), \\ h_{A,D}(x) &= 1, \quad x \in A, \\ h_{A,D}(x) &= 0, \quad x \in D. \end{aligned} \tag{3.5}$$

If we can show that the problem (3.3) has a *unique* solution, then we can be sure to have reduced the problem of computing probabilities $\mathbb{P}_x[\tau_A < \tau_D]$ to a problem of linear algebra.

Proposition 3.1. *Let Γ be a finite set, and let $A, D \subset \Gamma$ be non-empty. Assume that P is irreducible, i.e. that for any $x, y \in \Gamma$, there exists $n < \infty$ such that $p_n(x,y) > 0$. Then the problem (3.3) has a unique solution.*

The function $h_{A,D}$ is called the *equilibrium potential* of the *capacitor* A, B. The fact that

$$\mathbb{P}_x[\tau_A < \tau_D] = h_{A,D}(x) \tag{3.6}$$

for $x \in \Gamma \backslash (A \cup D)$ is the first fundamental relation between the theory of Markov chains and *potential theory.*

The next question is what happens for $x \in D$? Naturally, using the same reasoning as the one leading to (3.2), we obtain that

$$\begin{aligned} \mathbb{P}_x[\tau_A < \tau_D] = \sum_{y \in A} p(x,y) + \sum_{y \in \Gamma \backslash (A \cup D)} p(x,y) \mathbb{P}_y[\tau_A < \tau_D] \\ = \sum_{y \in \Gamma} p(x,y) h_{A,D}(y) \end{aligned} \tag{3.7}$$

It will be even more convenient to define, for all $x \in \Gamma$

$$e_{A,D}(x) \equiv -(Lh_{A,D})(x) \tag{3.8}$$

Then

$$\mathbb{P}_x[\tau_A < \tau_D] = \begin{cases} h_{A,D}(x), & \text{if } x \in \Gamma \backslash (A \cup D) \\ e_{A,D}(x), & \text{if } x \in D \\ 1 - e_{D,A}(x) & \text{if } x \in A \end{cases} \tag{3.9}$$

Let us now define the *capacity* of the capacitor A, D as

$$\text{cap}(A, D) \equiv \sum_{x \in D} \mu(x) e_{A,D}(x) \tag{3.10}$$

By the properties of $h_{A,D}$ it is easy to see that we can write

$$\begin{aligned} \sum_{x \in D} \mu(x) e_{A,D}(x) &= \sum_{x \in \Gamma} \mu(x)(1 - h_{A,D}(x))(-Lh_{A,,D})(x) \\ &= \sum_{x \in \Gamma} \mu(x) h_{A,D}(x)(Lh_{A,,D})(x) - \sum_{x \in \Gamma} \mu(x)(Lh_{A,,D})(x) \end{aligned} \tag{3.11}$$

Since $\mu(x)L = 0$, we get that

$$\text{cap}(A, D) = \sum_{x \in \Gamma} \mu(x) h_{A,D}(x)(Lh_{A,,D})(x) \equiv \Phi(h_{A,D}) \tag{3.12}$$

where

$$\Phi(h) \equiv \sum_{x \in \Gamma} \mu(x) h(x) Lh(x) = \frac{1}{2} \sum_{x,y} \mu(x) p(x, y) \left(h(x) - h(y)\right)^2 \tag{3.13}$$

is called the *Dirichlet form* associated to the Markov process with generator L. In fact, we will sometimes think of the Dirichlet form as the quadratic form associated to the generator and write

$$\Phi(f, g) \equiv (f, Lg)_\mu = \frac{1}{2} \sum_{x,y} \mu(x) p(x, y) \left(f(x) - f(y)\right) \left(g(x) - g(y)\right). \tag{3.14}$$

The representation of the capacity in terms of the Dirichlet form will turn out to be of fundamental importance. The reason for this is the ensuing variational representation, known as the *Dirichlet principle:*

Theorem 3.2. *Let $\mathcal{H}_D^A$ denote the space of functions*

$$\mathcal{H}_D^A \equiv \{h : \Gamma \to [0, 1], h(x) = 1, x \in A, h(x) = 0, x \in D\} \tag{3.15}$$

Then

$$\text{cap}(A, D) = \inf_{h \in \mathcal{H}_D^A} \Phi(h) \tag{3.16}$$

Moreover, the variational problem (3.15) has a unique minimizer that is given by the equilibrium potential $h_{A,D}$.

Proof. Differentiating $\Phi(h)$ with respect to $h(x)$ (for $x \in \Gamma\backslash(A \cup D)$) yields

$$\frac{\partial}{\partial h(x)}\Phi(h) = 2\mu(x)Lh(x) \tag{3.17}$$

Thus if h minimizes Φ, it must be true that $Lh(x) = 0$. Since we have already seen that the Dirichlet problem (3.3) has a unique solution, the theorem is proven.

While in general the capacity is a weighted sum over certain probabilities, if we choose for the set D just a point $x \in \Gamma$, we get that

$$\mathbb{P}_x[\tau_A < \tau_x] = \frac{1}{\mu(x)}\mathrm{cap}(A, x)$$

We will call these quantities sometimes *escape probabilities*. We see that they have, by virtue of Theorem 3.2 a direct variational representation. They play a crucial rôle in what will follow. Let us note the fact that $\mathrm{cap}(x, y) = \mathrm{cap}(y, x)$ implies that

$$\mu(x)\mathbb{P}_x[\tau_y < \tau_x] = \mu(y)\mathbb{P}_y[\tau_x < \tau_y] \tag{3.18}$$

which is sometimes helpful to get intuition. Note that this implies in particular that

$$\mathbb{P}_x[\tau_y < \tau_x] \leq \frac{\mu(y)}{\mu(x)}$$

which is quite often already a useful bound (provided of course $\mu(y) < \mu(x)$).

3.2 The One-Dimensional Chain

We will now consider the example of a one-dimensional nearest neighbor random walk (with inhomogeneous rates). For reasons that will become clear later, we introduce a parameter $\epsilon > 0$ and think of our state space as a one-dimensional "lattice" of spacing ϵ, that is we take $\Gamma \subset \epsilon\mathbb{Z}$, and transition probabilities

$$p(x, y) = \begin{cases} \sqrt{\frac{\mu(y)}{\mu(x)}}g(x, y), & \text{if } y = x \pm \epsilon, \\ 1 - p(x, x + \epsilon) - p(x, x - \epsilon), & \text{if } x = y, \\ 0, & \text{else} \end{cases} \tag{3.19}$$

where $\mu(x) > 0$, and g is such that $p(x, x) \geq 0$.

Equilibrium Potential.

Due to the one-dimensional nature of our process, we only equilibrium potentials we have to compute are of the form

$$h_{b,a}(x) = \mathbb{P}_x[\tau_b < \tau_a] \tag{3.20}$$

where $a < x < b$. The equations (3.5) then reduce to the one-dimensional discrete boundary value problem

$$\begin{aligned} p(x, x+\epsilon)(h(x+\epsilon)-h(x))+p(x, x-\epsilon)(h(x-\epsilon)-h(x)) &= 0, a < x < b \\ h(a) &= 0 \\ h(b) &= 1 \end{aligned}$$

We can solve this by recursion and get

$$h(x) = \frac{\sum_{y=a+\epsilon}^{x} \frac{1}{\mu(y)} \frac{1}{p(y,y-\epsilon)}}{\sum_{y=a+\epsilon}^{b} \frac{1}{\mu(y)} \frac{1}{p(y,y-\epsilon)}} \tag{3.21}$$

Capacities.

Given the explicit formula for the equilibrium potential, we can readily compute capacities. Without going into the detailed computations, I just quote the result:

$$\mathrm{cap}(a, b) = \frac{1}{\sum_{y=a+\epsilon}^{b} \frac{1}{\mu(y)} \frac{1}{p(y,y-\epsilon)}} \tag{3.22}$$

Remark 3.3. Formula (3.22) suggests another common "electrostatic" interpretation of capacities, namely as "resistances". In fact, if we interpret $\mu(x)p(x, x-\epsilon) = \mu(x-\epsilon)p(x-\epsilon, x)$ as the conductance of the "link" (resistor) $(x-\epsilon, x)$, then by Ohm's law, formula (3.22) represents the conductance of the chain of resistors from a to b. This interpretation is not restricted to the one-dimensional chain, but holds in general for reversible Markov chains. The capacity of the capacitor (A, D) may then be seen as the conductance of the resistor network between the two sets. In this context, the monotonicity properties of the capacities obtain a very natural interpretation: removing a resistor or reducing its conductivity can only decrease the conductivity of the network. There is a very nice account on the resistor network interpretation of Markov chains and some of its applications in a book by Doyle and Snell.

3.3 Mean Hitting Times

Our next task is to derive formulas for the mean values of hitting times τ_A. As in Section 3.1 we first derive a forward equation for $\mathbb{E}_x\tau_A$ by considering what can happen in the first step:

$$\mathbb{E}_x\tau_A = \sum_{y\in A} p(x,y) + \sum_{y\notin A} p(x,y)(1+\mathbb{E}_y\tau_A) \tag{3.23}$$

if $x \notin A$. If we define a function

$$w_A(x) \equiv \begin{cases} \mathbb{E}_x\tau_A, & \text{if } x \in \Gamma\backslash A \\ 0, & \text{if } x \in A \end{cases} \tag{3.24}$$

we see that (3.23) can be written in the nicer form

$$w_A(x) = \sum_{y\in\Gamma} p(x,y)w_A(y) + 1 \tag{3.25}$$

for $x \notin A$; i.e. w_A solves the *inhomogeneous Dirichlet problem*

$$\begin{aligned} Lw_A(x) &= 1, \quad x \in G\backslash A \\ w_A(x) &= 0, \quad x \in A \end{aligned} \tag{3.26}$$

Note that for $x \in A$ we can compute $\mathbb{E}_x\tau_A$ by considering the first step:

$$\mathbb{E}_x\tau_A = \sum_{y\in A} p(x,y) + \sum_{y\notin A} p(x,y)(1+\mathbb{E}_y\tau_A) \tag{3.27}$$

or in compact form

$$\mathbb{E}_x\tau_A = Pw_A(x) + 1 = -Lw_A(x) + 1 \tag{3.28}$$

Equations (3.26) is a special cases of the general Dirichlet problem

$$\begin{aligned} Lf(x) &= g(x), \quad x \in \Gamma\backslash A \\ f(x) &= 0, \quad x \in A \end{aligned} \tag{3.29}$$

for some set A and some function f. We have seen in Proposition 3.1 that the homogeneous boundary value problem (i.e. if $g \equiv 0$) has the unique solution $f(x) \equiv 0$. This implies that the problem (3.29) has a unique solution that can (by linearity) be represented in the form

$$f(x) = \sum_{y\in\Gamma\backslash A} G_{\Gamma\backslash A}(x,y)g(y) \tag{3.30}$$

Of course, $G_{\Gamma\setminus A}$ is simply the matrix inverse of the matrix $L^{\Gamma\setminus A}$ whose elements are

$$\begin{aligned} L^{\Gamma\setminus A}(x,y) &= L(x,y), \quad x,y \in \Gamma\setminus A \\ L^{\Gamma\setminus A}(x,y) &= 0, \quad x \in A \vee y \in A \end{aligned} \tag{3.31}$$

We will call $L^{\Gamma\setminus A}$ the Dirichlet operator on $\Gamma\setminus A$. Note that while L is a positive operator, due to Proposition 3.1, $L^{\Gamma\setminus A}$ is *strictly positive* whenever $A \neq \emptyset$. The inverse operator $G_{\Gamma\setminus A}(x,y)$ is usually called the *Green's function*.

We see that we would really like to compute this Green's function. What we will actually show now is that the Green's function can be computed in terms of equilibrium potentials and equilibrium measures. To see this, let us return to (3.8) and interpret this as an equation for $h_{D,A}$ where the boundary conditions are only prescribed on A but not on D: Note first that since $h_{A,D}(x) = 1 - h_{D,A}(x)$, (3.8) can also be written as

$$e_{A,D}(x) = Lh_{D,A}(x) \tag{3.32}$$

With this relation, assuming $e_{A,D}$ given on D, we can re-write the determining equation for $h_{D,A}$ as an inhomogeneous Dirichlet problem with boundary conditions only on A:

$$\begin{aligned} Lh_{D,A}(x) &= 0, \quad x \in \Gamma\setminus(A \cup D) \\ Lh_{D,A}(x) &= e_{A,D}(x), \quad x \in D \\ h_{D,A}(x) &= 0, \quad x \in A \end{aligned} \tag{3.33}$$

Thus we can write

$$h_{D,A}(x) = \sum_{y\in D} G_{\Gamma\setminus A}(x,y) e_{A,D}(y) \tag{3.34}$$

Let us now consider the special case when D is a single point, say $D = \{z\}$. Then (3.34) gives

$$h_{z,A}(x) = G_{\Gamma\setminus A}(x,z) e_{A,D}(z) \tag{3.35}$$

which gives immediately

$$G_{\Gamma\setminus A}(x,z) = \frac{h_{z,A}(x)}{e_{A,D}(z)}. \tag{3.36}$$

Now due to the symmetry of L,

$$G_{\Gamma\setminus A}(x,z)\mu(x) = G_{\Gamma\setminus A}(z,x)\mu(z) \tag{3.37}$$

This gives us

$$G_{\Gamma\setminus A}(z,x) = \frac{\mu(x)h_{z,A}(x)}{\mu(z)e_{A,D}(z)}. \tag{3.38}$$

In conclusion, we have the following useful fact:

Proposition 3.4. *The Dirichlet Green's function for any set $A \subset G$ can be represented in terms of the equilibrium potential and capacities as*

$$G_{\Gamma\setminus A}(x,z) = \frac{\mu(z)h_{x,A}(z)}{\mathrm{cap}(A,z)} \tag{3.39}$$

We now get immediately the desired representations for the mean times:

$$\mathbb{E}_x\tau_A = \sum_{y\in\Gamma\setminus A} \frac{\mu(y)h_{x,A}(y)}{\mathrm{cap}(A,x)} \tag{3.40}$$

These formulas will prove to be excessively useful in the sequel.

3.4 Renewal Equations

The application of Proposition 3.4 may not appear very convincing, as we can actually solve the Dirichlet problems directly. On the other hand, even if we admit that the Dirichlet variational principle gives us a good tool to compute the denominator, i.e. the capacity, we still do not know how to compute the equilibrium potential. We will now show that a surprisingly simple argument provides a tool that allows us to reduce, for our purposes, the computation of the equilibrium potential to that of capacities.

This yields the *renewal bound* for the equilibrium potential.

Lemma 3.5. *Let $A, D \subset \Gamma$ be disjoint, and $x \in (A\cup D)^c$. Then*

$$\mathbb{P}_x[\tau_A < \tau_D] = h_{A,D}(x) \le \frac{\mathrm{cap}(x,A)}{\mathrm{cap}(x,D)} \tag{3.41}$$

Proof. The basis of our argument is the trivial observation that if the process starting at a point x wants to realise the event $\{\tau_A < \tau_D\}$, it may do so by going to A immediately and without returning to x again, or it may return to x without either going to A or to D. Clearly, once the process returns to x it is in the same position as at the starting time, and we can use the (strong) Markov property to separate the

probability of what happened before the first return to x to whatever will happen later. Formally:

$$\begin{aligned}\mathbb{P}_x[\tau_A < \tau_D] &= \mathbb{P}_x[\tau_A < \tau_{D\cup x}] + \mathbb{P}_x[\tau_x < \tau_{A\cup D} \wedge \tau_A < \tau_D] \\ &= \mathbb{P}_x[\tau_A < \tau_{D\cup x}] + \mathbb{P}_x[\tau_x < \tau_{A\cup D}]\mathbb{P}_x[\tau_A < \tau_D] \quad (3.42)\end{aligned}$$

We call this a renewal equation. We can solve this equation for $\mathbb{P}_x[\tau_A < \tau_D]$:

$$\mathbb{P}_x[\tau_A < \tau_D] = \frac{\mathbb{P}_x[\tau_A < \tau_{D\cup x}]}{1 - \mathbb{P}_x[\tau_x < \tau_{A\cup D}]} = \frac{\mathbb{P}_x[\tau_A < \tau_{D\cup x}]}{\mathbb{P}_x[\tau_{A\cup D} < \tau_x]} \tag{3.43}$$

By elementary monotonicity properties this representation yields the bound

$$\mathbb{P}_x[\tau_A < \tau_D] \leq \frac{\mathbb{P}_x[\tau_A < \tau_x]}{\mathbb{P}_x[\tau_D < \tau_x]} = \frac{\mathrm{cap}(x, A)}{\mathrm{cap}(x, D)} \tag{3.44}$$

Of course this bound is useful only if $\frac{\mathrm{cap}(x,A)}{\mathrm{cap}(x,D)} < 1$, but since $\mathbb{P}_x[\tau_A < \tau_D] = 1 - \mathbb{P}_x[\tau_D < \tau_A]$, the applicability of this bound is quite wide. It is quite astonishing how far the simple use of this renewal bound will take us.

4 Metastability

We come now to a general definition of metastability in the context of discrete Markov chains.

4.1 Metastable Points

Definition 4.1. *Assume that Γ is a discrete set. Then a Markov processes X_t is metastable with respect to the set of points $\mathcal{M} \subset \Gamma$, if*

$$\frac{\sup_{x\in\mathcal{M}} \mathbb{P}_x[\tau_{\mathcal{M}\setminus x} < \tau_x]}{\inf_{y\notin\mathcal{M}} \mathbb{P}_y[\tau_{\mathcal{M}} < \tau_y]} \leq \rho \ll 1 \tag{4.1}$$

We will see that Definition 4.1 is (at least if Γ is finite) equivalent to an alternative definition involving averaged hitting times.

Definition 4.2. *Assume that Γ is a finite discrete set. Then a Markov processes X_t is metastable with respect to the set of points $\mathcal{M} \subset \Gamma$, if*

$$\frac{\inf_{x\in\mathcal{M}} \mathbb{E}_x\tau_{\mathcal{M}\setminus x}}{\sup_{y\notin\mathcal{M}} \mathbb{E}_y\tau_{\mathcal{M}}} \geq 1/\rho \gg 1 \tag{4.2}$$

We will show that without further assumptions on the particular properties of the Markov chain we consider, the fact that a set of metastable states satisfying the condition of Definition 4.1 exists implies a number of structural properties of the chain.

4.2 Ultrametricity

An important fact that allows to obtain general results under our Definition of metastability is the fact that it implies approximate ultrametricity of capacities. This has been noted in [8].

Lemma 4.3. *Assume that $x, y \in \Gamma$, $D \subset \Gamma$. Then, if for $0 < \delta < \frac{1}{2}$, $\mathrm{cap}(x, D) \leq \delta \mathrm{cap}(y, x)$, then*

$$\frac{1-2\delta}{1-\delta} \leq \frac{\mathrm{cap}(x,D)}{\mathrm{cap}(y,D)} \leq \frac{1}{1-\delta} \tag{4.3}$$

Proof. The key idea of the proof is to use the probabilistic representation of capacities and renewal type arguments involving the strong Markov property. It would be nice to have a purely analytic proof of this lemma.

We first prove the upper bound. We write

$$\mathrm{cap}(x, D) = \mathrm{cap}(D, x) = \sum_{z\in D} \mu(z) e_{x,D}(z) = \sum_{z\in D} \mu(z) \mathbb{P}_z[\tau_x < \tau_D] \tag{4.4}$$

Now

$$\begin{aligned} \mathbb{P}_z[\tau_x < \tau_D] &= \mathbb{P}_z[\tau_x < \tau_D, \tau_y < \tau_D] + \mathbb{P}_z[\tau_x < \tau_D, \tau_y \geq \tau_D] \\ &= \mathbb{P}_z[\tau_x < \tau_D, \tau_y < \tau_D] + \mathbb{P}_z[\tau_x < \tau_{D\cup y}] \mathbb{P}_x[\tau_D < \tau_y] \\ &= \mathbb{P}_z[\tau_x < \tau_D, \tau_y < \tau_D] + \mathbb{P}_z[\tau_x < \tau_{D\cup y}] \frac{\mathbb{P}_x[\tau_D < \tau_{y\cup x}]}{\mathbb{P}_x[\tau_{D\cup y} < \tau_x]} \end{aligned} \tag{4.5}$$

Here we used the Markov property at the optional time τ_x to split the second probability into a product, and then the renewal equation (3.43). Now by assumption,

$$\frac{\mathbb{P}_x[\tau_D < \tau_{y\cup x}]}{\mathbb{P}_x[\tau_{D\cup y} < \tau_x]} \leq \frac{\mathbb{P}_x[\tau_D < \tau_x]}{\mathbb{P}_x[\tau_y < \tau_x]} \leq \delta \tag{4.6}$$

Inserting (4.6) into (4.5) we arrive at

$$\begin{aligned} \mathbb{P}_z[\tau_x < \tau_D^x] &\leq \mathbb{P}_z[\tau_y < \tau_D, \tau_x < \tau_D] + \delta \mathbb{P}_z[\tau_x < \tau_{D\cup y}] \\ &\leq \mathbb{P}_z[\tau_y < \tau_D] + \delta \mathbb{P}_z[\tau_x < \tau_D] \end{aligned} \tag{4.7}$$

Inserting this inequality into (4.4) implies

$$\begin{aligned} \mathrm{cap}(x, D) &\leq \sum_{z\in D} \mu(z) \mathbb{P}_z[\tau_y < \tau_D] + \delta \sum_{z\in D} \mu(z) \mathbb{P}_z[\tau_x < \tau_D] \\ &= \mathrm{cap}(y, D) + \delta \mathrm{cap}(x, D) \end{aligned} \tag{4.8}$$

which implies the upper bound.

The lower bound follows by observing that from the upper bound we get that $\mathrm{cap}(x, D) \leq \frac{\delta}{1-\delta}\mathrm{cap}(x, y)$. Thus reversing the rôle of x and y, the resulting upper bound for $\frac{\mathrm{cap}(y,D)}{\mathrm{cap}(x,D)}$ is precisely the claimed lower bound.

Lemma 4.3 has the following immediate corollary, which is the version of the ultrametric triangle inequality we are looking for:

Corollary 4.4. *Let $x, y, z \in \mathcal{M}$. Then*

$$\mathrm{cap}(x, y) \geq \frac{1}{3} \min\left(\mathrm{cap}(x, z), \mathrm{cap}(y, z)\right) \tag{4.9}$$

Valleys. In the sequel it will be useful to have the notion of a "valley" or "attractor" of a point in $\mathcal{M}$. We set for $x \in \mathcal{M}$,

$$A(x) \equiv \left\{ z \in \Gamma \,|\, \mathbb{P}_z[\tau_x = \tau_{\mathcal{M}}] = \sup_{y \in \mathcal{M}} \mathbb{P}_z[\tau_y = \tau_{\mathcal{M}}] \right\} \tag{4.10}$$

Note that valleys may overlap, but from Lemma 4.3 it follows easily that the intersection has a vanishing invariant mass. The notion of a valley in the case of a process with invariant measure $\exp(-f(x)/\epsilon)$ coincides with this notion.

More precisely, the next Lemma will show that if y belongs to the valley of $m \in \mathcal{M}$, then either the capacity $\mathrm{cap}(y, \mathcal{M}\backslash m)$ is essentially the same as $\mathrm{cap}(m, \mathcal{M}\backslash m)$, or the invariant mass of y is excessively small. That is to say that within each valley there is a subset that "lies below the barrier defined by the capacity $\mathrm{cap}(m, \mathcal{M}\backslash m)$, while the rest has virtually no mass, i.e. the process never really gets there.

Lemma 4.5. *Let $m \in \mathcal{M}$, $y \in A(m)$, and $D \subset \mathcal{M}\backslash m$. Then either*

$$\frac{1}{2} \leq \frac{\mathrm{cap}(m, D)}{\mathrm{cap}(y, D)} \leq \frac{3}{2}$$

or

$$\mu(y) \leq 3|\mathcal{M}| \frac{\mu(y)}{\mathrm{cap}(y, \mathcal{M})} \mathrm{cap}(m, D)$$

Proof. Lemma 4.3 implies that *if* $\mathrm{cap}(m, y) \geq 3\mathrm{cap}(m, D)$, then (4.5) holds. Otherwise,

$$\frac{\mu(y)}{\mu(m)} \leq 3 \frac{\mu(y)}{\mathrm{cap}(y, m)} \frac{\mathrm{cap}(m, D)}{\mu(m)} \tag{4.11}$$

Since $y \in A(m)$, we have that $\mathbb{P}_y[\tau_m \leq \tau_{\mathcal{M}}] \geq 1/|\mathcal{M}|$. On the other hand, the renewal estimate yields

$$\mathbb{P}_y[\tau_m \leq \tau_{\mathcal{M}}] \leq \frac{\mathrm{cap}(y,m)}{\mathrm{cap}(y,\mathcal{M})} \tag{4.12}$$

Hence

$$\mathrm{cap}(y,\mathcal{M}) \leq |\mathcal{M}|\mathrm{cap}(y,m) \tag{4.13}$$

which yields (4.5).

4.3 Mean Hitting Times

We will now derive a very convenient expression for the mean time of arrival in a subset $J \subset \mathcal{M}$ of the metastable points. This will be based on our general representation formula for mean arrival times (3.40) together with the renewal based inequality for the equilibrium potential and the ultrametric inequalities for the capacities that we just derived under the hypothesis of Definition 4.1.

Let $x \in \mathcal{M}$, $x \notin J \subset \mathcal{M}$. We want to compute $\mathbb{E}_x\tau_J$. Our starting point is the following equation, that is immediate from (3.40)

$$\mathbb{E}_x\tau_J = \frac{\mu(x)}{\mathrm{cap}(x,J)} \sum_{y\in J^c} \frac{\mu(y)}{\mu(x)} h_{x,J\setminus x}(y) \tag{4.14}$$

We want to estimate the summands in the sum (4.14). We will set

$$a \equiv \inf_y \mu(y)^{-1}\mathrm{cap}(y,\mathcal{M}). \tag{4.15}$$

The following lemma provides the necessary control over the equilibrium potentials appearing in the sum.

Lemma 4.6. *Let $x \in \mathcal{M}$ and $J \subset \mathcal{M}$ with $x \notin J$. Then:*

(i) If $x = m$, either

$$h_{x,J}(y) \geq 1 - \frac{3}{2}|\mathcal{M}|a^{-1}\frac{\mathrm{cap}(x,J)}{\mu(y)} \tag{4.16}$$

or

$$\mu(y) \leq 3|\mathcal{M}|a^{-1}\mathrm{cap}(m,J) \tag{4.17}$$

(ii) If $m \in J$, then

$$\mu(y)h_{x,J}(y) \leq \frac{3}{2}|\mathcal{M}|a^{-1}\mathrm{cap}(m,x) \tag{4.18}$$

(iii) If $m \notin J \cup x$, then either

$$h_{x,J}(y) \leq 3\frac{\mathrm{cap}(m,x)}{\mathrm{cap}(m,J)} \tag{4.19}$$

and

$$h_{x,J}(y) \geq 1 - 3\frac{\mathrm{cap}(m,J)}{\mathrm{cap}(m,x)} \tag{4.20}$$

or

$$\mu(y) \leq 3|\mathcal{M}|a^{-1}\max\left(\mathrm{cap}(m,J), \mathrm{cap}(m,x)\right) \tag{4.21}$$

We will skip the somewhat tedious proof of this lemma. With its help one can give rather precise expressions for the mean hitting times (4.14) that only involve capacities and the invariant measure. We will only consider a special case of particular interest, namely when J contains all points in $\mathcal{M}$ that 'lie lower than' x, i.e. if $J = \mathcal{M}_x \equiv \{m \in \mathcal{M} : \mu(m) \geq \delta\mu(x)\}$, for some $\delta \ll 1$ to be chosen. We will call the corresponding time $\tau_{\mathcal{M}_x}$ the *metastable exit time* from x. In fact, it is reasonable to consider this the time when the process has definitely left x, since the mean time to return to x from $\mathcal{M}_x$ is definitely larger than (or at most equal in degenerate cases) $\mathbb{E}_x\tau_{\mathcal{M}_x}$. Nicely enough, these mean times can be computed very precisely:

Theorem 4.7. *Let $x \in \mathcal{M}$ and $J \subset \mathcal{M}\backslash x$ be such a that for all $m \notin J \cup x$ either $\mu(m) \ll \mu(x)$ or $\mathrm{cap}(m,J) \gg \mathrm{cap}(m,x)$, then*

$$\mathbb{E}_x\tau_J = \frac{\mu(A(x))}{\mathrm{cap}(x,J)}\left(1 + O(\rho)\right) \tag{4.22}$$

Proof. Left to the reader.

Finally we want to compute the mean time to reach $\mathcal{M}$ starting from a general point.

Lemma 4.8. *Let $z \notin \mathcal{M}$. Then*

$$\mathbb{E}_z\tau_{\mathcal{M}} \leq a^{-2}\left(|\{y : \mu(y) \geq \mu(z)|\} + C\right) \tag{4.23}$$

Proof. Using Lemma 4.2, we get that

$$\begin{aligned}
\mathbb{E}_z\tau_{\mathcal{M}} &\le \frac{\mu(z)}{\mathrm{cap}(z,\mathcal{M})}\sum_{y\in\mathcal{M}^c}\frac{\mu(y)}{\mu(z)}\max\left(1,\frac{\mathrm{cap}(y,z)}{\mathrm{cap}(y,\mathcal{M})}\right)\\
&= \frac{\mu(z)}{\mathrm{cap}(z,\mathcal{M})}\sum_{y\in\mathcal{M}^c}\frac{\mu(y)}{\mu(z)}\max\left(1,\frac{\mathbb{P}_y[\tau_z<\tau_y]}{\mathbb{P}_y[\tau_{\mathcal{M}}<\tau_y]}\right)\\
&\le \sup_{y\in\mathcal{M}^c}\left(\frac{\mu(y)}{\mathrm{cap}(y,\mathcal{M})}\right)^2\sum_{y\in\mathcal{M}^c}\max\left(\frac{\mu(y)}{\mu(z)},\mathbb{P}_z[\tau_y<\tau_z]\right)\\
&\le \sup_{y\in\mathcal{M}^c}\left(\frac{\mu(y)}{\mathrm{cap}(y,\mathcal{M})}\right)^2\left(\sum_{y:\mu(y)\le\mu(z)}\frac{\mu(y)}{\mu(z)}+\sum_{y:\mu(y)>\mu(z)}1\right)\\
&\le \sup_{y\in\mathcal{M}^c}\left(\frac{\mu(y)}{\mathrm{cap}(y,\mathcal{M})}\right)^2\left(C+|\{y:\mu(y)>\mu(z)\}|\right) \qquad (4.24)
\end{aligned}$$

which proves the lemma.

Remark 4.9. If Γ is finite (resp. not growing too fast with ϵ), the above estimate combined with Theorem 4.7 shows that the two definitions of metastability we have given in terms of mean times rep. capacities are equivalent. On the other hand, in the case of infinite state space Γ, we cannot expect the supremum over $\mathbb{E}_z\tau_{\mathcal{M}}$ to be finite, which shows that our second definition is less suitable than the first.

5 Upper and Lower Bounds for Capacities

In this lecture we will introduce some powerful, though simple ideas that allow to compute upper and lower bounds for capacities that are relevant for metastability. We will do this with a concrete model, the Glauber dynamics for the Curie-Weiss model, at hand, but the methods we will use are also applicable in other situations.

Let me therefore first of all recall this model and its dynamics.

5.1 The Curie-Weiss Model

The Curie-Weiss model is the simplest model for a ferromagnet. Here the state space is the hypercube $\mathcal{S}_N \equiv \{-1,1\}^N$, and the Hamiltonian of the Curie–Weiss model is

$$H_N(\sigma) = -\frac{1}{2N}\sum_{1\le i,j\le N}\sigma_i\sigma_j - h\sum_{i=1}^N\sigma_i \qquad (5.1)$$

The crucial feature of the model is that the Hamiltonian is a function of the *macroscopic variable*, the magnetization as a *function* on the configuration space: we will call

$$m_N(\sigma) \equiv N^{-1} \sum_{i=1}^{N} \sigma_i \tag{5.2}$$

the *empirical magnetization*. Here we divided by N to have a specific magnetization. A function of this type is called a *macroscopic* function, because it depends on all spin variables. We can indeed write

$$H_N(\sigma) = -\frac{N}{2} [m_N(\sigma)]^2 - hNm_N(\sigma) \equiv -NE_h(m_N(\sigma)) \tag{5.3}$$

The computation of the partition function is then very easy: We write

$$Z_{\beta,h,N} = \sum_{m \in \mathcal{M}_N} e^{N\beta(\frac{m^2}{2}+mh)} z_{m,N} \tag{5.4}$$

where $\mathcal{M}_N$ is the set of possible values of the magnetization, i.e.,

$$\begin{aligned} \mathcal{M}_N &\equiv \{m \in \mathbb{R} : \exists \sigma \in \{-1,1\}^N : m_N(\sigma) = m\} \\ &= \{-1, -1 + 2/N, \ldots, 1 - 2/N, 1\} \end{aligned} \tag{5.5}$$

and

$$z_{m,N} \equiv \sum_{\sigma \in \{-1,1\}^N} \mathbb{1}_{m_N(\sigma)=m} \tag{5.6}$$

is a 'micro-canonical partition function'. Fortunately, the computation of this micro-canonical partition function is easy. In fact, all possible values of m are of the form $m = 1 - 2k/N$, and for these

$$z_{m,N} = \binom{N}{N(1-m)/2} \equiv \frac{N!}{[N(1-m)/2]![N(1+m)/2]!} \tag{5.7}$$

It is always useful to know the asymptotics of the logarithm of the binomial coefficients. If we set, for $m \in \mathcal{M}_N$

$$N^{-1} \ln z_{m,N} \equiv \ln 2 - I_N(m) \equiv \ln 2 - I(m) - J_N(m) \tag{5.8}$$

where

$$I(m) = \frac{1+m}{2} \ln(1+m) + \frac{1-m}{2} \ln(1-m) \tag{5.9}$$

then

$$J_N(m) = \frac{1}{2N}\ln\frac{1-m^2}{4} + \frac{\ln N + \ln(2\pi)}{2N} + O\Big(N^{-2}\Big(\frac{1}{1-m}+\frac{1}{1+m}\Big)\Big) \tag{5.10}$$

(5.10) is obtained using the asymptotic expansion for the logarithm of the Gamma function. The function $I(x)$ is called *Cramèr's entropy function* and worth memorizing. Note that by its nature it is a relative entropy. The function J_N is of lesser importance, since it is very small.

The Gibbs measure is then

$$\mu_{\beta,N} \equiv \frac{\exp\big(\beta N\big[m_N(\sigma)^2/2 + hm_N(\sigma)\big]\big)}{Z_{\beta,N}}. \tag{5.11}$$

an important role is played by the measure induced by the map m_N,

$$\mathcal{Q}_{\beta,N}(m) \equiv \mu_{\beta,N}\circ m_N^{-1}(m) = \frac{\exp\left(-\beta N\left[-E_h(m)\right] - NI_N(m)\right)}{2^N Z_{\beta,N}}. \tag{5.12}$$

Note that this measure concentrates sharply, as N goes to infinity, on the minimizers of the function $F_{\beta,N} \equiv -E_h(m) + \beta^{-1}I(m)$.

5.2 Glauber Dynamics

Typical dynamics studied for such models are Glauber dynamics, i.e. (random) Markov chains $\sigma(t)$, defined on the configuration space $\mathcal{S}^N$ that are reversible with respect to the (random) Gibbs measures $\mu_{\beta,N}(\sigma)$ and in which the transition rates are non-zero only if the final configuration can be obtained from the initial one by changing the value of one spin only. A particular choice of transition rates are given by the Metropolis algorithm:

$$p_N(\sigma,\sigma') \equiv \begin{cases} 0, & \text{if } \|\sigma-\sigma'\| > 2,\\ \frac{1}{N}e^{-\beta[H_N(\sigma')-H_N(\sigma)]_+}, & \text{if } \|\sigma-\sigma'\| = 2,\\ 1-\frac{1}{N}\sum_{\tau:\|\tau-\sigma\|=2} e^{-\beta[H_N(\sigma')-H_N(\sigma)]_+}, & \text{if } \sigma=\sigma'. \end{cases} \tag{5.13}$$

Here $[f]_+ \equiv \max(f,0)$.

There is a simple way of analysing this dynamics which is based on the observation that in this particular model, if $\sigma(t)$ is the Markov process with the above transition rates, then the stochastic process

$\tilde{m}_N(t) \equiv m_N(\sigma(t))$ is again a Markov process with state space $\mathcal{M}_N$ and invariant measure $\mathcal{Q}_{\beta,N}$.

Here we do not want to follow this course, but we will use more generally applicable bounds that will, however, reproduce the exact results in this simple case.

As a first problem that we encounter in this way it the proper definition of metastable state. Since the invariant (Gibbs) measure is constant on the sets of configurations with given value of m_N, clearly looking for configurations that are local minima of the energy, H_N, is not a good idea. In fact, since the induced measure $\mathcal{Q}_{\beta,N}$ has local maxima at the minima of the function $f_{\beta,N}$, and given the symmetries of the problem, it seems far more natural to consider as metastable sets the sets

$$M_\pm \equiv \{\sigma : m_N(\sigma) = m_\pm^*\}, \tag{5.14}$$

where $m_\pm^*$ are the largest, respectively smallest local minimizer of $f_{\beta,N}(m) = 0$.

We may come back to the question whether this is a feasible definition later. For the moment, we want to see how in such a situation we can compute the relevant capacity, $\mathrm{cap}(M_+, M_-)$.

5.3 Upper Bounds

Our task is to compute

$$\mathrm{cap}(M_+, M_-) = \inf_{h\in\mathcal{H}} \frac{1}{2} \sum_{\sigma,\tau\in\mathcal{S}_N} \mu(\sigma) p_N(\sigma,\tau) \left[h(\sigma) - h(\tau)\right]^2, \tag{5.15}$$

where

$$\mathcal{H} = \{h : \Sigma_N \to [0,1] : h(\sigma) = 0, \sigma \in M_+, h(\sigma) = 1, \sigma \in M_-\}. \tag{5.16}$$

The general strategy is to prove an upper bound by guessing some a-priori properties of the minimizer, h, and then to find the minimizers within this class. There are no limits to one's imagination here, but of course some good physical insight will be helpful. The good thing is that, whatever we will guess here, will be put to the test later when we will or will not be able to come up with a matching lower bound. Quite often it is not a bad idea to try to assume that the minimizer (i.e. the equilibrium potential) depends on σ only through some *order parameter*. In our case this can only be the magnetisation, $m_N(\sigma)$. As a matter of fact, due to symmetry, in our case we can know a priori

that this will be true for a fact, but, even if it may not be true, it may give a good bound for the capacity: it is really only necessary that this assumption holds in those places where the sum in (5.15) gives a serious contribution!

Let us see where this gets us:

$$\mathrm{cap}(M_+, M_-) = \inf_{g\in\widetilde{\mathcal{H}}} \frac{1}{2} \sum_{\sigma,\tau\in\mathcal{S}_N} \mu(\sigma)p_N(\sigma,\tau)\left[g(m_N(\sigma)) - g(m_N(\tau))\right]^2, \tag{5.17}$$

where

$$\widetilde{\mathcal{H}} = \left\{g : [-1,1] \to [0,1] : g(m_-^*) = 0, g(m_+^*) = 1\right\}. \tag{5.18}$$

But

$$\begin{aligned}
&\frac{1}{2} \sum_{\sigma,\tau\in\mathcal{S}_N} \mu(\sigma)p_N(\sigma,\tau)\left[g(m_N(\sigma)) - g(m_N(\tau))\right]^2 \qquad (5.19)\\
&= \frac{1}{2} \sum_{m,m'} [g(m) - g(m')]^2 \sum_{\sigma:m_N(\sigma)=m,\tau:m_N(\tau)=m'} \mu(\sigma)p_N(\sigma,\tau)\\
&= \frac{1}{2} \sum_{m,m'} \mathcal{Q}_{\beta,N}(m) r_N(m.m')[g(m) - g(m')]^2,
\end{aligned}$$

where

$$r_N(x,y) \equiv \frac{1}{\mathcal{Q}_{\beta,N}(x)} \sum_{\sigma:m_n(\sigma)=x} \sum_{\tau:m_N(\tau)=y} \mu_{\beta,N}(\sigma)p_N(\sigma,\tau) \tag{5.20}$$

In our special case of the Metropolis dynamics, $p_N(\sigma,\tau)$ depends only on $m_N(\sigma)$ and $m_N(\tau)$

$$r_N(x,y) = \begin{cases} 0, & \text{if } |x-y| > 2/N,\\ (1-x)/2\exp(-\beta N|F_N(x+2/N) - F_N(x)]_+, & \text{if } y = x+2/N,\\ (1+x)/2\exp(-\beta N|F_N(x-2/N) - F_N(x)]_+, & \text{if } y = x-2/N,\\ 1 - \frac{(1-x)}{2}\exp(-\beta N|F_N(x+2/N) - F_N(x)]_+ & \\ -\frac{(1+x)}{2}\exp(-\beta N|F_N(x-2/N) - F_N(x)]_+, & \text{if } x = y. \end{cases} \tag{5.21}$$

The main point is that the remaining one-dimensional variational problem involving the quadratic form (5.19) can be solved exactly. The answer is given in the form

$$\inf_{g\in\tilde{\mathcal{H}}} \frac{1}{2}\sum_{m,m'} \mathcal{Q}_{\beta,N} r_N(m.m')[g(m)-g(m')]^2 \tag{5.22}$$

$$= \left[\sum_{\ell=0}^{N(m_+-m_-)/2-1} \frac{1}{\mathcal{Q}_{\beta,N}(m_+ 2\ell/N) r_N(m_- + 2\ell/N, m_-(2\ell+2)/N)}\right]^{-1}$$

The sum appearing in the denominator can be further analysed using the Laplace method, but this shall be not our main concern at the moment.

The question we want to address now is how to get a corresponding lower bound.

5.4 Lower Bounds

The real art in analysing metastability in our approach lies in the judicious derivation of lower bounds for the capacity. There are two ways of seeing how this can be done. First, we may use the monotonicity of the Dirichlet form in the parameters $p_N(\sigma,\tau)$. This means that we may, in particular, set a number of the $p_N(\sigma,\tau)$ to zero to obtain a simpler system for which we may be able to find the solution of our variational problem more easily. In many cases, this strategy has provided good results.

There is, however, a more general approach that gives us far more flexibility. To this end, consider a countable set I, and a let $\mathcal{G} \equiv \{g_{xy}, x, y \in \Gamma\}$, be a collection of sub-probability measures on I, i.e. for each (x,y), $g_{xy}(\alpha) \geq 0$, and $\sum_{\alpha\in I} g_{xy}(\alpha) \leq 1$. Then

$$\begin{aligned}
\mathrm{cap}(A,D) &= \inf_{h\in\mathcal{H}_{A,D}} \sum_{\alpha\in I} \frac{1}{2}\sum_{x,y} \mu(y) g_{xy}(\alpha) p(x,y) \left\|h(x)-h(y)\right\|^2 \\
&\geq \sum_{\alpha\in I} \inf_{h\in\mathcal{H}_{A,D}} \frac{1}{2}\sum_{x,y} \mu(y) g_{xy}(\alpha) p(x,y) \left\|h(x)-h(y)\right\|^2 \\
&\equiv \sum_{\alpha\in I} \inf_{h\in\mathcal{H}_{A,D}} \Phi^{\mathcal{G}(\alpha)}(h) \equiv \sum_{\alpha\in I} \mathrm{cap}^{\mathcal{G}(\alpha)}(A,D),
\end{aligned} \tag{5.23}$$

where $\mathcal{H}_{A,D}$ is the space of functions from Γ to $[0,1]$ that vanish on D and are equal to one on A. As this it true for all $\mathcal{G}$, we get the variational principle

$$\mathrm{cap}(A,D) = \sup_{\mathcal{G}} \sum_{\alpha\in I} .\mathrm{cap}^{\mathcal{G}(a)}(A,D) \tag{5.24}$$

Note that this may look trivial, as of course the supremum is realised for the trivial case $I = \{1\}$, $g_{xy}(1) = 1$, for all (x, y). The interest in the principle arises from the fact that there may be other choices that still realise the supremum (or at least come very close to it). If we denote by $h_{A,D}^{\mathcal{G}(\alpha)}$ the minimizer of $\Phi^{\mathcal{G}(\alpha)}(h)$, then $\mathcal{G}$ realises the supremum, whenever

$$h_{A,D}^{\mathcal{G}(\alpha)}(x) = h_{A,D}(x), \quad \forall x : g_{xy}(\alpha) \neq 0. \tag{5.25}$$

Of course we do not know $h_{A,D}(x)$, but this observation suggest a very good strategy to prove lower bounds, anyhow: guess a plausible test function h for the upper bound, then try to construct $\mathcal{G}$ such that the minimizers, $h^{\mathcal{G}(\alpha)}$, are computable, and are similar to h! If this succeeds, the resulting upper and lower bounds will be at least very close. Remarkably, this strategy actually does work in many cases.

Lower Bounds Through One-Dimensional Paths

The following approach was developed in this context with D. Ioffe [5]. It can be seen as a specialisation of a more general approach by Berman and Konsowa [2]. We describe it first in an abstract context and then apply it to the Curie-Weiss model. Let $\Gamma \equiv \Gamma_0 \cup \ldots \Gamma_K$ be the vertex set of a graph. We call a graph layered, if for any edge, $e \equiv (v, u)$, there exists ℓ such that $u \in \Gamma_\ell$ and $v \in \Gamma_{\ell-1}$ or $v \in \Gamma_{\ell+1}$. Let $p(u, v)$ be a Markov transition matrix whose associated graph is a layered graph on Γ, and whose unique reversible measure is given by μ. We are interested in computing the capacity from Γ_0 to Γ_K, i.e.

$$\begin{aligned} C_{0,K} \equiv & \frac{1}{2} \inf_{h:h(\Gamma_0)=1,h(\Gamma_K)=0} \sum_{\sigma,\sigma' \in \Gamma} \mu(\sigma) p(\sigma, \sigma') \left[h(\sigma) - h(\sigma')\right]^2 \qquad (5.26) \\ = & \inf_{h:h(\Gamma_0)=1,h(\Gamma_K)=0} \sum_{\ell=0}^{K-1} \sum_{\sigma_\ell \in \Gamma_\ell, \sigma_{\ell+1} \in \Gamma_{\ell+1}} \mu(\sigma_\ell) p(\sigma_\ell, \sigma_{\ell+1}) \left[h(\sigma_\ell) - h(\sigma_{\ell+1})\right]^2 \end{aligned}$$

Let us introduce a probability measure ν_0 on Γ_0. Let q be a Markov transition matrix on Γ whose elements, $q(\sigma.\sigma')$, are non-zero only if, for some ℓ, $\sigma \in \Gamma_\ell$ and $\sigma' \in \Gamma_{\ell+1}$, and if $p(\sigma, \sigma') > 0$. Define, for $\ell \geq 0$,

$$\nu_{\ell+1}(\sigma_{\ell+1}) = \sum_{\sigma_\ell \in \Gamma_\ell} \nu_\ell(\sigma_\ell) q(\sigma_\ell, \sigma_{\ell+1}). \tag{5.27}$$

Let $\mathcal{T}$ denote the set of all directed paths form Γ_0 to Γ_K on our graph. Note that the Markov chain with transition matrix q and initial distribution ν_0 defines a probability measure on $\mathcal{T}$, which we will denote by $\mathbb{Q}$.

We now associate for any $T \in \mathcal{T}$ and any edge, $b = (\sigma_\ell, \sigma_{\ell+e})$ in our graph the weight

$$w_T(b) \equiv \begin{cases} 0, & \text{if } b \notin T \\ \mathbb{Q}(T)/(q(b)\nu_\ell(\sigma_\ell)), & \text{if } b = (\sigma_\ell, \sigma_{\ell+1}) \in T \end{cases} \tag{5.28}$$

Lemma 5.1. *For all b in our graph,*

$$\sum_T w_T(b) = 1 \tag{5.29}$$

Proof. Note that, if $T = (\sigma - 1, \ldots, \sigma_K)$, and $b = (\sigma_\ell, \sigma_{\ell+e})$

$$\begin{aligned} \mathbb{Q}(T)/(q(b)\nu_\ell(\sigma_\ell)) = \nu_0(\sigma_0)q(\sigma_0, \sigma_1)\ldots \\ \ldots q(\sigma_{\ell-1}, \sigma_\ell)\frac{1}{\nu_\ell}q(\sigma_{\ell+1}, \sigma_{\ell+2})\ldots q(\sigma_{k-1}, \sigma_K) \end{aligned} \tag{5.30}$$

Summing over all T containing b means to sum this expression over $\sigma_0, \sigma_1, \ldots, \sigma_{\ell-1}$, and over $\sigma_{\ell+1}, \ldots, \sigma_K$. Using the definition of ν_k is easy to see that this gives exactly one.

Theorem 5.2. *With the definition above we have that*

$$C_{0,K} \geq \sum_{T\in\mathcal{T}} \mathbb{Q}(T) \left[\sum_{\ell=0}^{K-1} \frac{\nu_\ell(\sigma_\ell)q(\sigma_\ell, \sigma_{\ell+1})}{\mu(\sigma_\ell)p(\sigma_\ell, \sigma_{\ell+1})}\right]^{-1} \tag{5.31}$$

Proof. In view of the preceding lemma, we have clearly that

$$\begin{aligned} C_{0,K} &= \inf_{h:h(\Gamma_0)=1, h(\Gamma_K)=0} \sum_{\ell=0}^{K-1} \sum_{\sigma_\ell\in\Gamma_\ell, \sigma_{\ell+1}\in\Gamma_{\ell+1}} \sum_{T\in\mathcal{T}} w_T(\sigma_\ell, \sigma_{\ell+1})\mu(\sigma_\ell) \\ &\quad \times p(\sigma_\ell, \sigma_{\ell+1})\left[h(\sigma_\ell) - h(\sigma_{\ell+1})\right]^2 \\ &= \inf_{h:h(\Gamma_0)=1, h(\Gamma_K)=0} \sum_{T\in\mathcal{T}} \mathbb{Q}(T) \sum_{\ell=0}^{K-1} \frac{\mu(\sigma_\ell)p(\sigma_\ell, \sigma_{\ell+1})}{\nu_\ell(\sigma_\ell)q(\sigma_\ell, \sigma_{\ell+1})}\left[h(\sigma_\ell) - h(\sigma_{\ell+1})\right]^2 \\ &\geq \sum_{T\in\mathcal{T}} \mathbb{Q}(T) \inf_{h:h(\sigma_0)=1, h(\sigma_K)=0} \sum_{\ell=0}^{K-1} \frac{\mu(\sigma_\ell)p(\sigma_\ell, \sigma_{\ell+1})}{\nu_\ell(\sigma_\ell)q(\sigma_\ell, \sigma_{\ell+1})}\left[h(\sigma_\ell) - h(\sigma_{\ell+1})\right]^2 \end{aligned} \tag{5.32}$$

Solving the one-dimensional variational problems in the last line gives the well-known expression that is given in the statement of the theorem.

Remark 5.3. The quality of the lower bound depends on to what extend the interchange of the summation over paths and the infimum over the functions h is introducing errors. If the minimizers are the same for all paths, then no error what so ever is made. This will be the case if the effective capacities

$$\frac{\mu(\sigma_\ell)p(\sigma_\ell,\sigma_{\ell+1})}{\nu_\ell(\sigma_\ell)q(\sigma_\ell,\sigma_{\ell+1})}$$

are independent of the particular path.

Remark 5.4. Berman and Konsowa [2] prove a more general lower bound where the space of paths contains all self-avoiding paths, without the restriction of directedness we have made. In this class, they show the supremum over all probability distributions on the space of paths yields exactly the capacity.

Application to the Curie-Weiss Model

In the Curie-Weiss model, it is a very simple matter to achieve the objective stated in the remark above. Clearly, we choose for the layers the sets $\Gamma_\ell \equiv \{\sigma : m_N(\sigma) = m_-^* + 2\ell/N\}$.

Since $\mu(\sigma)$ depends only on $m_N(\sigma)$, and $p_N(\sigma,\tau)$ depends only on $m_N(\sigma)$, $m_N(\tau)$, and the fact whether or not τ is reachable from σ by a single spin flip, it is enough to choose for ν_0 the uniform measure on the set Γ_0, and for $q(\sigma_\ell,\sigma_{\ell+1}) = \frac{2}{N-Nm_-^*-2\ell}$. Then ν_ℓ is the uniform measure on Γ_ℓ, and that

$$\frac{\nu_\ell(\sigma_\ell)}{\mu(\sigma_\ell)} = \frac{1}{\mu(\Gamma_\ell)} = \frac{1}{\mathcal{Q}_{\beta,N}(m_-^* + 2\ell/N)}, \tag{5.33}$$

and

$$\frac{p_N(\sigma_\ell,\sigma_{\ell+1})}{q(\sigma_\ell,\sigma_{\ell+1})} = r_N(\sigma_\ell,\sigma_{\ell+1}). \tag{5.34}$$

Thus, the lower bound from Theorem 5.2 reproduces the upper bound exactly.

6 Metastability and Spectral Theory

We now turn to the characterisation of metastability through spectral data. The connection between metastable behaviour and the existence of small eigenvalues of the generator of the Markov process has been

realised for a very long time. Some key references are [16, 17, 18, 23, 24, 27, 31, 33, 38, 40, 39]. Here we will explain the approach developed in [8].

We will show that Definition 4.1 implies that the spectrum of L decomposes into a cluster of $|\mathcal{M}|$ very small real eigenvalues that are separated by a gap from the rest of the spectrum. To avoid complications we will assume that $|\Gamma|$ s finite throughout this section.

6.1 Basic Notions

Let $D \subset \Gamma$. We say that $\lambda \in \mathbb{C}$ is an *eigenvalue* for the Dirichlet problem, resp. the Dirichlet operator L^D, with boundary conditions in D if the equation

$$\begin{aligned} Lf(x) &= \lambda f(x), \quad x \in \Gamma \backslash D \\ f(x) &= 0, \quad x \in D \end{aligned} \tag{6.1}$$

has a non-zero solution f. $f \equiv f_\lambda$ is then called an eigenfunction. If $D = \emptyset$ we call the corresponding values eigenvalues of L. From the symmetry of the operator L it follows that any eigenvalue must be real; moreover, since L is positive, all eigenvalues are positive. If Γ is finite and $D \neq \emptyset$, the eigenvalues of the corresponding Dirichlet problem are *strictly* positive, while zero is an eigenvalue of L itself with the constant function the corresponding (right) eigenfunction.

If λ is not an eigenvalue of L^D, the Dirichlet problem

$$\begin{aligned} (L - \lambda) f(x) &= g(x), \quad x \in \Gamma \backslash D \\ f(x) &= 0, \quad x \in D \end{aligned} \tag{6.2}$$

has a unique solution and the solution can be represented in the form

$$f(x) = \sum_{y \in \Gamma \backslash D} G^\lambda_{\Gamma \backslash D}(x, y) g(y) \tag{6.3}$$

where $G^\lambda_{\Gamma \backslash D}(x, y)$ is called the Dirichlet Green's function for $L - \lambda$.

Equally, the boundary value problem

$$\begin{aligned} (L - \lambda) f(x) &= 0, \quad x \in \Gamma \backslash D \\ f(x) &= \phi(x), \quad x \in D \end{aligned} \tag{6.4}$$

has a unique solution in this case. Of particular importance will be the λ-equilibrium potential (of the capacitor (A, D)), $h^\lambda_{A,D}$, defined as the solution of the Dirichlet problem

$$(L-\lambda)h^{\lambda}_{A,D}(x) = 0, \quad x \in (A\cup D)^c$$
$$h^{\lambda}_{A,D}(x) = 1, \quad x \in A$$
$$h^{\lambda}_{A,D}(x) = 0, \quad x \in D \tag{6.5}$$

We may define analogously the λ-equilibrium measure

$$e^{\lambda}_{D,A}(x) \equiv (L-\lambda)h^{\lambda}_{A,D}(x) \tag{6.6}$$

Alternatively, $e^{\lambda}_{A,D}$ on A, is the unique measure on A, such that

$$h^{\lambda}_{A,D}(x) = \sum_{y\in A} G^{\lambda}_{D^c}(x,y)e^{\lambda}_{A,D}(y) \tag{6.7}$$

If $\lambda \neq 0$, the equilibrium potential still has a probabilistic interpretation in terms of the Laplace transform of the hitting time τ_A of the process starting in x and killed in D. Namely, we have for general λ, that, with $u(\lambda) \equiv -\ln(1-\lambda)$,

$$h^{\lambda}_{A,D}(x) = \mathbb{E}_x e^{u(\lambda)\tau_A}\mathbb{1}_{\tau_A<\tau_D}$$

for $x \in (A\cup D)^c$, whenever the right-hand side exists. Note that the left hand side is in general the meromorphic extension (in $\lambda \in \mathbb{C}$) of the probabilistically defined right-hand side.

6.2 A Priori Estimates

The first step of our analysis consists in showing that the matrix $L^{\mathcal{M}}$ that has Dirichlet conditions in all the points of $\mathcal{M}$ has a minimal eigenvalue that is not smaller than $O(a^2)$.

The basis for a priori estimates of eigenvalues is the variational representation of the principal eigenvalue:

Lemma 6.1. *The principal (smallest) eigenvalue, λ_D, of the Dirichlet operator L^D satisfies*

$$\lambda_D = \inf_{f:f(x)=0,x\in D} \frac{\Phi(f)}{\|f\|^2_{2,\mu}} \tag{6.8}$$

where $\|f\|_{2,\mu} \equiv \left(\sum_{x\in\Gamma}\mu(x)f(x)^2\right)^{1/2}$

Proof. Since L^D is a positive operator, there exists A such that $L = A^*A$. If λ is the smallest eigenvalue of L^D, then $\sqrt{\lambda}$ is the smallest eigenvalue of A and vice versa. But

$$\begin{aligned}\lambda &= \left(\inf_{f:f(x)=0,x\in D} \frac{\|Af\|_{2,\mu}}{\|f\|_{2,\mu}}\right)^2 \\ &= \inf_{f:f(x)=0,x\in D} \frac{\|Af\|_{2,\mu}^2}{\|f\|_{2,\mu}^2} = \inf_{f:f(x)=0,x\in D} \frac{\Phi(f)}{\|f\|_{2,\mu}^2} \end{aligned} \tag{6.9}$$

The following is a simple application due to Donsker and Varadhan [15].

Lemma 6.2. *Let λ_D denote the infimum of the spectrum of L^D. Then*

$$\lambda_D \geq \frac{1}{\sup_{z\in\Gamma\setminus D} \mathbb{E}_z \tau_D} \tag{6.10}$$

Proof. Consider any function $\phi : \Gamma \to \mathbb{R}$ satisfying $\phi(x) = 0$ for $x \in \Delta$. We will use the elementary fact that for all $x, y \in \Gamma$ and $C > 0$

$$\phi(y)\phi(x) \leq \frac{1}{2}(\phi(x)^2 C + \phi(y)^2/C) \tag{6.11}$$

with $C \equiv \psi(y)/\psi(x)$, for some positive function ψ to get a lower bound on $\Phi(\phi)$:

$$\begin{aligned}\Phi(\phi) &= \frac{1}{2}\sum_{x,y} \mu(x)p(x,y)\left(\phi(x)-\phi(y)\right)^2 \\ &= \|\phi\|_{2,\mu}^2 - \sum_{x,y\notin D} \mu(x)p(x,y)\phi(x)\phi(y) \\ &\geq \|\phi\|_{2,\mu}^2 - \sum_{x,y} \mu(x)p(x,y)\frac{1}{2}\left(\phi(x)^2\psi(y)/\psi(x) + \phi(y)^2\psi(x)/\psi(y)\right) \\ &= \|\phi\|_{2,\mu}^2 - \sum_{x\notin D} \mu(x)\phi(x)^2 \frac{\sum_y p(x,y)\psi(y)}{\psi(x)} \end{aligned} \tag{6.12}$$

Now choose $\psi(x) = w_D(x)$ (defined in (3.24)). By (3.25), this yields

$$\begin{aligned}&\Phi(\phi) \geq \|\phi\|_{2,\mu}^2 - \|\phi\|_{2,\mu}^2 + \sum_{x\notin D} \mu(x)\phi(x)^2 \frac{1}{w_D(x)} \\ &= \sum_{x\notin D} \mu(x)\phi(x)^2 \frac{1}{w_D(x)} \geq \|\phi\|_{2,\mu}^2 \sup_{x\in D^c} \frac{1}{w_D(x)} = \|\phi\|_{2,\mu}^2 \frac{1}{\sup_{x\in D^c} \mathbb{E}_x \tau_D} \end{aligned} \tag{6.13}$$

Since this holds for all ϕ that vanish on D,

$$\lambda_D = \inf_{\phi:\phi(x)=0,x\in D} \frac{\Phi(\phi)}{\|\phi\|^2_{2,\mu}} \geq \frac{1}{\sup_{x\in D^c} \mathbb{E}_x \tau_D} \tag{6.14}$$

as claimed. ere a is defined in (4.15).

If we combine this result with the estimate from Lemma 4.8, we obtain the following proposition.

Proposition 6.3. *Let λ^0 denote the principal eigenvalue of the operator $L^{\mathcal{M}}$. Then there exists a constant $C > 0$, independent of ϵ, such that for all ϵ small enough,*

$$\lambda^0 \geq Ca^2, \tag{6.15}$$

where a is defined in (4.15).

Remark 6.4. Proposition 6.3 links the fast time scale to the smallest eigenvalue of the Dirichlet operator, as should be expected. Note that the relation is not very precise. We will soon derive a much more precise relation between times and eigenvalues for the cluster of small eigenvalues.

6.3 Characterization of Small Eigenvalues

We will now obtain a representation formula for all eigenvalues that are smaller than λ^0. It is clear that there will be precisely $|\mathcal{M}|$ such eigenvalues. This representation was exploited in [8], but already in 1973 Wentzell [40, 39] put forward very similar ideas (in the case of general Markov processes). As will become clear, this is extremely simple in the context of discrete processes (see [10] for the more difficult continuous case).

The basic idea is to use the fact that the solution of the Dirichlet problem

$$\begin{aligned} (L-\lambda)f(x) &= 0, \quad x \notin \mathcal{M} \\ f(x) &= \phi_x, \quad x \in \mathcal{M}, \end{aligned} \tag{6.16}$$

which exists uniquely if $\lambda < \lambda^0$, already solves the eigenvalue equation $L\phi(x) = \lambda\phi(x)$ everywhere, except possibly on $\mathcal{M}$. It is natural to try to choose the boundary conditions ϕ_x, $x \in \mathcal{M}$ carefully in such a way that $(L-\lambda)f(x) = 0$ holds also for all $x \in \mathcal{M}$. Note that there are $|\mathcal{M}|$ free parameters $(\phi_x,\, x \in \mathcal{M})$ for just as many equations. Moreover, by linearity,

$$f(y) = \sum_{x \in \mathcal{M}} \phi_x h^{\lambda}_{x,\mathcal{M}\setminus x}(y). \tag{6.17}$$

Thus the system of equations to be solved can be written as

$$0 = \sum_{x \in \mathcal{M}} \phi_x L h^{\lambda}_{x,\mathcal{M}\setminus x}(m) \equiv \sum_{x \in \mathcal{M}} \phi_x e^{\lambda}_{x,\mathcal{M}\setminus x}(m), \quad \forall\, m \in \mathcal{M}, \tag{6.18}$$

Thus, if these equations have a non-zero solution $\phi_x x \in \mathcal{M}$, then λ is an eigenvalue. On the other hand, if λ is an eigenvalue smaller than λ^0 with eigenfunction ϕ_λ, then we may take $\phi_x \equiv \phi_\lambda(x)$ in (6.16). Then, obviously, $f(y) = \phi_\lambda(y)$ solves (6.16) uniquely, and it must be true that (6.18) has a non-zero solution.

Let us denote by $\mathcal{E}_{\mathcal{M}}(\lambda)$ the $|\mathcal{M}| \times |\mathcal{M}|$- matrix with elements

$$(\mathcal{E}_{\mathcal{M}}(\lambda))_{xy} \equiv e^{\lambda}_{z,\mathcal{M}\setminus z}(x). \tag{6.19}$$

Since the condition for (6.16) to have a non-zero solution is precisely the vanishing of the determinant of $\mathcal{E}^{\lambda}_{\mathcal{M}}$, we can now conclude that:

Lemma 6.5. *A number $\lambda < \lambda^0$ is an eigenvalue of L if and only if*

$$\det \mathcal{E}_{\mathcal{M}}(\lambda) = 0 \tag{6.20}$$

In the following we need a useful expression for the matrix elements of $\mathcal{E}_{\mathcal{M}}(\lambda)$. Since we anticipate that λ will be small, we set

$$h^{\lambda}_x(y) \equiv h_x(y) + \psi^{\lambda}_x(y), \tag{6.21}$$

where $h_x(y) \equiv h_{x,\mathcal{M}\setminus x}(y)$ and consequently $\psi^{\lambda}_x(y)$ solves the inhomogeneous Dirichlet problem

$$\begin{aligned} (L-\lambda)\psi^{\lambda}_x(y) &= \lambda h_x(y), \quad y \in \Gamma\setminus\mathcal{M} \\ \psi^{\lambda}_x(y) &= 0, \quad y \in \mathcal{M} \end{aligned} \tag{6.22}$$

A reorganisation of terms allows to express the matrix $\mathcal{E}_{\mathcal{M}}(\lambda)$ in the following form:

Lemma 6.6.

$$(\mathcal{E}_{\mathcal{M}}(\lambda))_{xz} = \mu(x)^{-1}\Big(\Phi(h_z, h_x) - \lambda((h_z, h_x)_\mu + (h_x, \psi^{\lambda}_z)_\mu\Big) \tag{6.23}$$

Proof. Note that

$$(L-\lambda)h_z^\lambda(x) = (L-\lambda)h_z(x) + (L-\lambda)\psi_z^\lambda(x) \\ = Lh_z(x) - \lambda h_z(x) + (L-\lambda)\psi_z^\lambda(x) \quad (6.24)$$

Now,

$$Lh_z(x) = \frac{\mu(x)}{\mu(x)} h_x(x) Lh_z(x) \quad (6.25)$$

The function $\mu^{-1}(y')h_x(y')Lh_z(y')$ vanishes for all $y' \neq x$. Thus, by adding a huge zero,

$$Lh_z(x) = \mu(x)^{-1} \sum_{y' \in \Gamma} \mu(y') h_x(y') Lh_z(y') \\ = \mu(x)^{-1} \frac{1}{2} \sum_{y,y' \in \Gamma} \mu(y') p(y',y))[h_z(y') - h_z(y)][h_x(y') - h_x(y)] \quad (6.26)$$

there the second inequality is obtained just as in the derivation of the representation of the capacity through the Dirichlet form. Similarly,

$$(L-\lambda)\psi_z^\lambda(x) = \\ \mu(x)^{-1} \sum_{y' \in \Gamma} \mu(y') \left(h_x(y')(L-\lambda)\psi_z^\lambda(y') - \lambda \mathbb{1}_{y' \neq x} h_x(y') h_z(y') \right) \quad (6.27)$$

Since $\psi_z^\lambda(y) = 0$ whenever $y \in \mathcal{M}$, and $Lh_x(y)$ vanishes whenever $y \notin \mathcal{M}$, using the symmetry of L, we get that the right-hand side of (6.27) is equal to

$$-\lambda \mu(x)^{-1} \sum_{y' \in \Gamma} \left(\mu(y') h_x(y') (\psi_z^\lambda(y') + \mathbb{1}_{y' \neq x} h_x(y') h_z(y') \right) \quad (6.28)$$

Adding the left-over term $-\lambda h_z(x) = -\lambda h_x(x) h_z(x)$ from (6.24) to (6.27), we arrive at (6.23).

Expanding in λ

Anticipating that we are interested in small λ, we want to control the λ-dependent terms ψ^λ in the formula for the matrix $\mathcal{E}_{\mathcal{M}}(\Lambda)$. From (6.22) we can conclude immediately that ψ_x^λ is small compared to h_x in the $L_2(\Gamma, \mu)$ sense when λ is small, since

$$\psi_x^\lambda = \lambda (L^{\mathcal{M}} - \lambda)^{-1} h_x. \quad (6.29)$$

Using that for symmetric operators, $\|(L-a)^{-1}\| \leq \frac{1}{\mathrm{dist(\,spec(L),a)}}$, we see that

$$\|\psi_x^\lambda\|_{2,\mu} \leq \frac{\lambda}{\lambda^0 - \lambda}\|h_x\|_{2,\mu}. \tag{6.30}$$

We are now in a position to relate the small eigenvalues of L to the eigenvalues of the classical capacity matrix. Let us denote by $\|\cdot|_2 \equiv \|\cdot\|_{2,\mu}$.

Theorem 6.7. *If $\lambda < \lambda^0$ is an eigenvalue of L, then there exists an eigenvalue μ of the $|\mathcal{M}| \times |\mathcal{M}|$-matrix $\mathcal{K}$ whose matrix elements are given by*

$$\begin{aligned}\mathcal{K}_{zx} &= \frac{\frac{1}{2}\sum_{y\neq y'}\mu(y')p(y',y)[h_z(y')-h_z(y)][h_x(y')-h_x(y)]}{\|h_z\|_2\|h_x\|_2} \\ &\equiv \frac{\Phi(h_z,h_x)}{\|h_z\|_2\|h_x)\|_2}\end{aligned} \tag{6.31}$$

such that $\lambda = \mu\,(1+O(\rho))$, where $\rho = \lambda/\lambda_0$.

We will skip the proof of this theorem since it is not really needed. In fact we will prove the following theorem.

Theorem 6.8. *Assume that there exists $x \in \mathcal{M}$ such that, for some $\delta \ll 1$*

$$\delta^2 \frac{\mathrm{cap}(x,\mathcal{M}\backslash x)}{\|h_x\|_2^2} \geq \max_{z\in\mathcal{M}\backslash x}\frac{\mathrm{cap}(z,\mathcal{M}\backslash z)}{\|h_z\|_2^2}. \tag{6.32}$$

Then the largest eigenvalue of L below λ_0 is given by

$$\lambda_x = \frac{\mathrm{cap}(x,\mathcal{M}\backslash x)}{\|h_x\|_2^2}(1+O(\delta^2+\rho^2)). \tag{6.33}$$

Moreover, the eigenvector, ϕ, corresponding to the largest eigenvalues normalized s.t. $\phi_x = 1$ satisfies $\phi_z \leq C(\delta+\rho)$, for $z \neq x$.

Proof. Let x be the point in $\mathcal{M}$ specified in the hypothesis. Denote by $\bar{\lambda}_1$ the Dirichlet eigenvalue with respect the set $\mathcal{M}\backslash x$. It is not very hard to verify that $\bar{\lambda}_1 \sim \frac{\mathrm{cap}(x,\mathcal{M}\backslash x)}{\|h_x\|_2^2}$. Moreover, one can easily verify that there will be exactly $|\mathcal{M}|-1$ eigenvalues below $\bar{\lambda}_1$. Thus, there must be one eigenvalue, λ_x, between $\bar{\lambda}_1$ and λ_0. We are trying to compute the precise value of this one, i.e. we look for a root of the determinant of $\mathcal{E}_M(\lambda)$ that is of order at least $\frac{\mathrm{cap}(x,\mathcal{M}\backslash x)}{\|h_x\|_2^2}$.

The determinant of $\mathcal{E}_{\mathcal{M}}(\lambda)$ vanishes together with that of the matrix $\widehat{\mathcal{K}}$ whose elements are

$$\widehat{\mathcal{K}}_{xz}=\frac{\mu(x)}{\|h_x\|_2\|h_z\|_2}\left(\mathcal{E}_{\mathcal{M}}(\lambda)\right)_{xz}=\frac{\Phi(h_x,h_z)}{\|h_x\|_2\|h_z\|_2}-\lambda\left(\frac{(h_x,h_z)_\mu+(\psi^\lambda_x,h_z)_\mu}{\|h_x\|_2\|h_z\|_2}\right). \tag{6.34}$$

We will now control al the elements of this matrix. We first deal with the off-diagonal elements of this matrix.

Lemma 6.9. *There is a constant $C<\infty$ such that*

$$\max_{x\neq z\in\mathcal{M}}\frac{(h_x,h_z)_\mu}{\|h_x\|_2\|h_z\|_2}\leq Ca^{-1}\max_{m\in\mathcal{M}}\mu(m)^{-1}\mathrm{cap}(m,\mathcal{M}\backslash m)\leq C\rho. \tag{6.35}$$

Proof. Note first by the estimate (3.44) the equilibrium potentials $h_x(y)$ are essentially equal to one on $A(x)$. Namely,

$$1\geq h_x(y)\geq 1-\frac{\mathrm{cap}(y,\mathcal{M}\backslash x)}{\mathrm{cap}(y,x)} \tag{6.36}$$

By Corollary 4.4, $\mathrm{cap}(y,\mathcal{M}\backslash x)\leq 2\mathrm{cap}(x,\mathcal{M}_x)$, or $\mu(y)\leq\frac{3|\mathcal{M}|}{a}\mathrm{cap}(x,\mathcal{M}_x)$.

Thus

$$\begin{aligned}\sum_{y\in A(x)}\mu(y)h_x(y)^2&\geq\sum_{\substack{y\in A(x)\\ \mu(y)\geq\frac{3|\mathcal{M}|}{a}\mathrm{cap}(x,\mathcal{M}_x)}}\mu(y)\left(1-\frac{\mathrm{cap}(x,\mathcal{M}\backslash x)}{\mathrm{cap}(y,x)}\right)^2\\&\geq\sum_{\substack{y\in A(x)\\ \mu(y)\geq\frac{3|\mathcal{M}|}{a}\mathrm{cap}(x,\mathcal{M}_x)}}\mu(y)-\sum_{y\in A(x)}2\frac{\mu(y)}{\mathrm{cap}(y,x)}\mathrm{cap}(x,\mathcal{M}\backslash x)\\&=\mu(A(m))\left(1-3|A(m)||\mathcal{M}|a^{-1}\frac{\mathrm{cap}(x,\mathcal{M}\backslash x)}{\mu(A(m))}\right)\\&\geq\mu(A(m))\left(1-O(\rho)\right).\end{aligned} \tag{6.37}$$

Thus the denominator in (6.35) is bounded from below by

$$\sqrt{\sum_{y\in A(x)}\mu(y)h_x^2(y)\sum_{y\in A(y)}\mu(y)h_z^2(y)}\geq\sqrt{\mu(A(x))\mu(A(z))}(1-O(\rho)). \tag{6.38}$$

To bound the numerator, we use that, for any $x\neq z\in\mathcal{M}$,

$$\sum_{y\in\Gamma}\mu(y)h_x(y)h_z(y)\leq C\rho\sqrt{\mu(x)\mu(z)}. \tag{6.39}$$

Using this bound we arrive at the assertion of the lemma.

Next we bound the terms involving ψ^λ.

Lemma 6.10. *If λ^0 denotes the principal eigenvalue of the operator L with Dirichlet boundary conditions in $\mathcal{M}$, then*

$$\left|\sum_{y\in\Gamma}\mu(y)\left(h_z(y)\psi_x^\lambda(y)\right)\right| \leq \frac{\lambda}{(\lambda^0-\lambda)}\|h_z\|_2\|h_x\|_2. \tag{6.40}$$

Proof. Recall that ψ_x^λ solves the Dirichlet problem (6.22). But the Dirichlet operator $L^{\mathcal{M}}-\lambda$ is invertible for $\lambda<\lambda^0$ and is bounded as an operator on $\ell^2(\Gamma,\mu)$ by $1/(\lambda^0-\lambda)$. Thus

$$\|\psi_x^\lambda|_2^2 \leq \left(\frac{\lambda}{\lambda^0-\lambda}\right)^2\|h_x\|_2^2 \tag{6.41}$$

The assertion of the lemma now follows from the Cauchy-Schwartz inequality.

Finally we come to the control of the terms involving $\Phi(h_x,h_z)$. By the Cauchy-Schwartz inequality,

$$\begin{aligned}\Phi(h_z,h_x) &= \left|\frac{1}{2}\sum_{y,y'}\mu(y')p(y',y)[h_x(y')-h_x(y)][h_z(y')-h_z(y)]\right| \\ &\leq \sqrt{\Phi(h_x)\Phi(h_z)}.\end{aligned} \tag{6.42}$$

Thus

$$\left|\frac{\Phi(h_x,h_z)}{\|h_x\|_2\|h_z\|_2}\right| \leq \sqrt{\frac{\Phi(h_x)}{\|h_x\|_2^2}}\sqrt{\frac{\Phi(h_z)}{\|h_z\|_2^2}}. \tag{6.43}$$

Therefore, by assumption, there exists one $x\in\mathcal{M}$ such that for any $(z,y)\neq(x,x)$,

$$\left|\frac{\Phi(h_x,h_z)}{\|h_x\|_2\|h_z\|_2}\right| \leq \delta\frac{\Phi(h_x)}{\|h_x\|_2^2}. \tag{6.44}$$

If we collect all our results:

(i) The matrix $\widetilde{\mathcal{K}}$ has one diagonal element

$$\widetilde{\mathcal{K}}_{xx} = \frac{\Phi(h_x)}{\|h_x\|_2^2} - \lambda(1+O(\lambda)) \equiv A - \lambda(1+O(\lambda)), \tag{6.45}$$

(ii) all other diagonal elements, $\mathcal{K}_{yy}$, satisfy

$$\widetilde{\mathcal{K}}_{yy} = O(\delta^2)A - \lambda(1+O(\lambda)) \approx -\lambda. \tag{6.46}$$

(iii) All off-diagonal elements satisfy

$$|\widetilde{\mathcal{K}}_{yz}| \leq C\delta^2 \frac{\Phi(h_x)}{\|h_x\|_2^2} + C\lambda\rho \equiv C(\delta A + \lambda\rho). \tag{6.47}$$

One can now look for non-zero solutions of the equations

$$\sum_y \widetilde{\mathcal{K}}_{zy} c_y = 0, \quad z \in \mathcal{M}. \tag{6.48}$$

In the sequel C denotes a numerical constant whose value changes from line to line. We may choose the vector c in such a way that $\max_{y\in\mathcal{M}} |c_y| = 1$, and this component realising the maximum to be equal to $+1$. We will first show that $c_x = 1$. To do so, assume that $c_z = 1$ for $z \neq x$. Then the equation (6.48) can be written

$$-\widetilde{\mathcal{K}}_{zz} = \sum_{y\neq z} c_y \widetilde{\mathcal{K}}_{zy} \tag{6.49}$$

Using our bounds, this implies

$$\lambda \leq C(\delta A + \rho\lambda) \Rightarrow \lambda \leq \frac{C\delta A}{1 - C\rho}, \tag{6.50}$$

in contradiction with the fact that $\lambda \geq A$. Thus $c_x = 1 \geq |c_z|$, for all $z \neq x$. Let us return to equation (6.48) for $z \neq x$. It now reads

$$-\widetilde{\mathcal{K}}_{zz} c_z = \sum_{y\neq z} c_y \widetilde{\mathcal{K}}_{zy}, \tag{6.51}$$

and hence

$$|c_z| \leq C \frac{\delta A + \rho\lambda}{\lambda} \tag{6.52}$$

Finally, we consider equation (6.48) with $z = x$,

$$\widetilde{\mathcal{K}}_{xx} = \sum_{y\neq x} c_y \widetilde{\mathcal{K}}_{xy}. \tag{6.53}$$

In view of our bounds on $\widetilde{\mathcal{K}}_{xy}$ and on c_y, this yields

$$|\widetilde{\mathcal{K}}_{xx}| \leq C \frac{(\delta A + \rho\lambda)^2}{\lambda} \leq C\delta^2 A + C\rho^2\lambda, \tag{6.54}$$

that is, we obtain that

$$|A - \lambda| \leq C\delta^2 A + \rho^2\lambda \tag{6.55}$$

which implies

$$\lambda = A\left(1 + O(\delta^2 + \rho^2)\right), \tag{6.56}$$

which is the first claim of the proposition. The assertion on the eigenvector follows from our estimates on the vector c.

Theorem 6.8 has the following simple corollary, that allows in many situations a complete characterization of the small eigenvalues of L.

Theorem 6.11. *Assume that we can construct a sequence of metastable sets $\mathcal{M}_k \supset \mathcal{M}_{k-1} \supset \cdots \supset \mathcal{M}_2 \supset \mathcal{M}_1 = x_0$, such that for any i, $\mathcal{M}_i \backslash \mathcal{M}_{i-1} = x_i$ is a single point, and that each $\mathcal{M}_i$ satisfies the assumptions of Theorem 6.8. Then L has k eigenvalues*

$$\lambda_i = \frac{\mathrm{cap}(x_i, \mathcal{M}_{i-1})}{\mu(A(x_i))}\,(1 + O(\delta)) \tag{6.57}$$

As a consequence,

$$\lambda_i = \frac{1}{\mathbb{E}_{x_i}\tau_{\mathcal{M}_{x_i}}}(1 + O(\delta)) \tag{6.58}$$

The corresponding normalized eigenfunction is given by

$$\psi_i(y) = \frac{h_{x_i,\mathcal{M}_{i-1}}(y)}{\|h_{x_i,\mathcal{M}_{i-1}}\|_2} + \sum_{j=1}^{i-1} O(\delta)\frac{h_{x_i,\mathcal{M}_{j-1}}(y)}{\|h_{x_i,\mathcal{M}_{j-1}}\|_2} \tag{6.59}$$

Proof. The idea behind this theorem is simple. Let the sets $\mathcal{M}_i$ of the corollary be given by $\mathcal{M}_i = \{x_1, \ldots, x_i\}$. Having computed the largest eigenvalue, λ_k, of L, we only have to search for eigenvalues smaller than λ_k. If we could be sure that the principal Dirichlet eigenvalue $\Lambda_{\mathcal{M}_{k-1}}$ is (much) larger than $k-1$st eigenvalue of L, then we could do so as before but replacing the set $\mathcal{M} \equiv \mathcal{M}_k$ by $\mathcal{M}_{k-1}$ everywhere. λ_{k-1} would then again be the largest eigenvalue of a capacity matrix involving only the points in $\mathcal{M}_{k-1}$. Iterating this procedure we arrive at the conclusion of the theorem.

The theorem is now immediate except for the statement (6.58). To conclude, we need to show that $\mathrm{cap}(x_{\ell+1}, \mathcal{M}_\ell) = \mathrm{cap}(x_\ell, \mathcal{M}_{x_\ell})$. To see this, note first that $\mathcal{M}_\ell \supset \mathcal{M}_{x_\ell}$. For if there was $x \in \mathcal{M}_{x_\ell}$ that is not contained in $\mathcal{M}_\ell$, then $\mathrm{cap}(x, \mathcal{M}_\ell \backslash x) \sim \mathrm{cap}(x_{\ell+1}, \mathcal{M}_\ell)$, while $\|h_{x_{\ell+1},\mathcal{M}_\ell}\|_2 \leq \|h_{x,\mathcal{M}_{\ell+1}\backslash x}\|_2$, contradicting the assumption in the construction of the set $\mathcal{M}_\ell$. Thus $\mathrm{cap}(x_{\ell+1}, \mathcal{M}_\ell) \geq \mathrm{cap}(x_\ell, \mathcal{M}_{x_\ell})$.

Similarly, if there was any point $x \in \mathcal{M}_\ell$ for which $\mathrm{cap}(x_{\ell+1}, \mathcal{M}_\ell) < \mathrm{cap}(x_\ell, \mathcal{M}_{x_\ell})$, then this point would have been associated to a larger eigenvalue in an earlier stage of the construction and thus would have already been removed from $\mathcal{M}_{\ell+1}$ before $x_{\ell+1}$ is being removed.

This observation allows us to finally realize that the k smallest eigenvalues of L are precisely the inverses of the mean (metastable) exit times from the metastable points $\mathcal{M}$.

6.4 Exponential Law of the Exit Time

The spectral estimates can be used to show that the law of the metastable exit times are close to exponential, provided the non-degeneracy hypothesis of Theorem 6.8 hold. Note that

$$\mathbb{P}_x[\tau_{\mathcal{M}_x} > t] = \sum_{x_1,\dots,x_t \notin \mathcal{M}_x} p(x,x_1) \prod_{i=1}^{t-1} p(x_i,x_{i+1}) = \sum_{y \notin \mathcal{M}_x} \left(P^{\mathcal{M}_x}\right)^t_{xy}. \tag{6.60}$$

To avoid complications, let us assume that the P is positive (in particular that P has no eigenvalues close to -1. This can be avoided e.g. by imposing that $p(x,x) > 0$). We now introduce the projection operators $\mathit{\Pi}$ on the eigenspace of the principal eigenvalue of $P^{\mathcal{M}_\ell}$. Then

$$\left(P^{\mathcal{M}_x}\right)^t_{xy} = \sum_{y \notin \mathcal{M}_x} \left(\left(P^{\mathcal{M}_x}\right)^t \mathit{\Pi}\right)_{xy} + \sum_{y \notin \mathcal{M}_x} \left(\left(P^{\mathcal{M}_x}\right)^t \mathit{\Pi}^c\right)_{xy}. \tag{6.61}$$

Using our estimate for the principal eigenfunction of $L^{\mathcal{M}_x}$ the first term in (6.61) equals

$$\left(1-\lambda^{\mathcal{M}_x}\right)^t \sum_{y \notin \mathcal{M}_x} \frac{h_{x,\mathcal{M}_x}(y)}{\|h_{x,\mathcal{M}_x}(y)\|_2}(1+O(\lambda^{\mathcal{M}_x})) \sim e^{-\lambda^{\mathcal{M}_x} t}. \tag{6.62}$$

The remaining term is bounded in turn by

$$e^{-\lambda_2^{\mathcal{M}_x} t}, \tag{6.63}$$

which under our assumptions decays much faster to zero than the first term.

References

1. G. Ben Arous and R. Cerf. Metastability of the three-dimensional Ising model on a torus at very low temperatures. *Electron. J. Probab.*, 1:no. 10 (electronic), 1996.
2. K. A. Berman and M. H. Konsowa. Random paths and cuts, electrical networks, and reversible Markov chains. *SIAM J. Discrete Math.*, 3:311–319, 1990.

3. A. Bovier. Metastability: a potential theoretic approach In *(International Congress of Mathematicians. Vol. III)*, pages 499–518, Eur. Math. Soc., Zürich, 2006.
4. A. Bovier. Metastability and ageing in stochastic dynamics. In *Dynamics and randomness II, Nonlinear Phenom. Complex Systems)*, pages 17–79, Kluwer Acad. Publ., Dordrecht, 2004.
5. A. Bianchi, A. Bovier, and D. Ioffe. Sharp estimates for metastable exists in the random field curie-weiss model. WIAS-preprint, 2008.
6. A. Bovier, F. den Hollander, and F.R. Nardi. Sharp asymptotics for kawasaki dynamics on a finite box with open boundary conditions. *Probab. Theor. Rel. Fields.*, 135:265–310, 2006.
7. A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in stochastic dynamics of disordered mean-field models. *Probab. Theor. Rel. Fields*, 119:99–161, 2001.
8. A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability and low-lying spectra in reversible Markov chains. *Commun. Math. Phys.*, 228:219–255, 2002.
9. A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes i. sharp asymptotics for capacities and exit times. *J. Europ. Math. Soc. (JEMS)*, 6:399–424, 2004.
10. A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes ii. precise asymptotics for small eigenvalues. *J. Europ. Math. Soc. (JEMS)*, 7:69–99, 2005.
11. A. Bovier and F. Manzo. Metastability in Glauber dynamics in the low-temperature limit: beyond exponential asymptotics. *J. Statist. Phys.*, 107:757–779, 2002.
12. M. Cassandro, A. Galves, E. Olivieri, and M. E. Vares. Metastable behavior of stochastic dynamics: a path-wise approach. *J. Statist. Phys.*, 35(5-6):603–634, 1984.
13. O. Catoni and R. Cerf. The exit path of a Markov chain with rare transitions. *ESAIM Probab. Statist.*, 1:95–144 (electronic), 1995/97.
14. R. Cerf. A new genetic algorithm. *Ann. Appl. Probab.*, 6(3):778–817, 1996.
15. D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes.* Springer Series in Statistics. Springer-Verlag, New York, 1988.
16. E.B. Davies. Metastable states of symmetric Markov semi-groups. I. *Proc. Lond. Math. Soc. III, Ser.*, 45:133–150, 1982.
17. E.B. Davies. Metastable states of symmetric Markov semi-groups. II. *J. Lond. Math. Soc. II, Ser.*, 26:541–556, 1982.
18. E.B. Davies. Spectral properties of metastable Markov semi-groups. *J. Funct. Anal.*, 52:315–329, 1983.
19. M. V. Day. On the exit law from saddle points. *Stochastic Process. Appl.*, 60(2):287–311, 1995.
20. W.Th.F. den Hollander. Three lectures on metastability under stochastic dynamics. In *this volume.* Springer, Berlin, 2008.

21. H. Eyring. The activated complex in chemical reactions. *J. Chem. Phys.*, 3:107–115, 1935.
22. L. R. Fontes, P. Mathieu, and P. Picco. On the averaged dynamics of the random field Curie-Weiss model. *Ann. Appl. Probab.*, 10(4):1212–1245, 2000.
23. M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1998.
24. B. Gaveau and L. S. Schulman. Theory of non-equilibrium first-order phase transitions for stochastic dynamics. *J. Math. Phys.*, 39(3):1517–1533, 1998.
25. B. Helffer, M. Klein, and F. Nier. Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach. *Mat. Contemp.*, 26:41–85, 2004.
26. B. Helffer and F. Nier. *Hypo-elliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, volume 1862 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.
27. R. A. Holley, S. Kusuoka, and D. W. Stroock. Asymptotics of the spectral gap with applications to the theory of simulated annealing. *J. Funct. Anal.*, 83(2):333–347, 1989.
28. Y. Kifer. Random perturbations of dynamical systems: a new approach. In *Mathematics of random media (Blacksburg, VA, 1989)*, volume 27 of *Lectures in Appl. Math.*, pages 163–173. Amer. Math. Soc., Providence, RI, 1991.
29. Vassili N. Kolokoltsov. *Semiclassical analysis for diffusions and stochastic processes*, volume 1724 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.
30. H.A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.
31. P. Mathieu. Spectra, exit times and long times asymptotics in the zero white noise limit. *Stoch. Stoch. Rep.*, 55:1–20, 1995.
32. P. Mathieu and P. Picco. Metastability and convergence to equilibrium for the random field Curie-Weiss model. *J. Statist. Phys.*, 91(3-4):679–732, 1998.
33. L. Miclo. Comportement de spectres d'opérateurs de Schrödinger à basse température. *Bull. Sci. Math.*, 119(6):529–553, 1995.
34. F. Nier. Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach. In *Journées "Équations aux Dérivées Partielles"*, pages Exp. No. VIII, 17. École Polytech., Palaiseau, 2004.
35. E. Olivieri and E. Scoppola. Markov chains with exponentially small transition probabilities: first exit problem from a general domain. I. The reversible case. *J. Statist. Phys.*, 79(3-4):613–647, 1995.

36. E. Olivieri and E. Scoppola. Markov chains with exponentially small transition probabilities: first exit problem from a general domain. II. The general case. *J. Statist. Phys.*, 84(5-6):987–1041, 1996.
37. R. H. Schonmann and S. B. Shlosman. Wulff droplets and the metastable relaxation of kinetic Ising models. *Comm. Math. Phys.*, 194(2):389–462, 1998.
38. E. Scoppola. Renormalization and graph methods for Markov chains. In *Advances in dynamical systems and quantum physics (Capri, 1993)*, pages 260–281. World Sci. Publishing, River Edge, NJ, 1995.
39. A. D. Ventcel′. The asymptotic behavior of the largest eigenvalue of a second order elliptic differential operator with a small parameter multiplying the highest derivatives. *Dokl. Akad. Nauk SSSR*, 202:19–22, 1972.
40. A. D. Ventcel′. Formulas for eigenfunctions and eigenmeasures that are connected with a Markov process. *Teor. Verojatnost. i Primenen.*, 18:3–29, 1973.

Three Lectures on Metastability under Stochastic Dynamics

Frank den Hollander[1,2]

[1] Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands
denholla@math.leidenuniv.nl
[2] EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Contents

1 Introduction

Metastability is a phenomenon where a physical, chemical or biological system, under the influence of a noisy dynamics, moves between different regions of its state space on different time scales. On short

R. Kotecký (ed.), *Methods of Contemporary Mathematical Statistical Physics*, Lecture Notes in Mathematics 1970, DOI: 10.1007/978-3-540-92796-9_5,

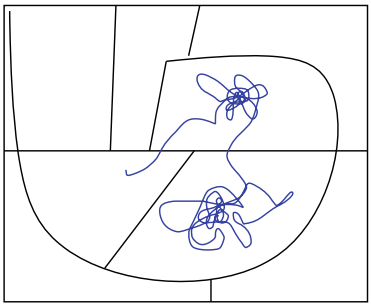


Fig. 1. The paradigm picture of metastability.

time scales the system is in a *quasi-equilibrium* within a single region, while on long time scales it undergoes *rapid transitions between quasi-equilibria* in different regions (see Fig. 1).

Examples of metastability can be found in:

- *biology*: folding of proteins;
- *climatology*: effects of global warming;
- *economics*: crashes of financial markets;
- *materials science*: anomalous relaxation in disordered media;
- *physics*: freezing of supercooled liquids.

The task of mathematics is to formulate *microscopic* models of the relevant underlying dynamics, to prove the occurrence of metastable behavior in these models on *macroscopic* space-time scales, and to identify the key mechanisms behind the experimentally observed *universality* in the metastable behavior of whole classes of systems. This is a challenging program!

The mathematics of metastability started around 1935, with the work of Eyring, Kramers and Wigner on diffusions in potential wells. It further developed in the 1970's, through the work of Lebowitz and Penrose on metastable states in van der Waals theory [25] and Freidlin and Wentzell on randomly perturbed dynamical systems [15]. It accelerated in the 1980's with the implementation of Freidlin-Wentzell theory in statistical physics by Capocaccia, Cassandro, Galves, Kotecký, Martinelli, Neves, Olivieri, Schonmann, Scoppola and Vares. Presently, metastability is a highly active subfield of probability theory and statistical physics.

Two approaches to metastability are central within mathematics:

- *Pathwise approach*: This was initiated in 1984 by Cassandro, Galves, Olivieri and Vares [11], and is based on monitoring the *full trajectory* of the dynamics, in the spirit of Freidlin-Wentzell theory.

- *Potential-theoretic approach*: This was initiated in 2001 by Bovier, Eckhoff, Gayrard and Klein [5], [6], and is based on an electric network perspective of the dynamics, focussing on crossing times via estimates on *capacities*.

The latter approach is highlighted in the paper by Bovier in the present volume [4]. For recent overviews of metastability, see the monograph by Olivieri and Vares [24] and the review papers by den Hollander [17] and Bovier [3]. Earlier review papers include Penrose and Lebowitz [25], Schonmann [28], [29], Scoppola [31], Vares [32], Olivieri and Scoppola [23].

In Lectures 1–3 below we describe the metastable behavior of Ising spins subject to Glauber dynamics and of lattice gas particles subject to Kawasaki dynamics, both in two dimensions. Attention focusses on the identification of the geometry of the critical droplet for the crossover from the metastable state to the stable state, and on the estimation of the crossover time. We consider three cases:

(1) finite systems at low temperature;
(2) large systems at low temperature;
(3) moderate systems at positive temperature.

These cases are progressively more challenging, and for the latter two work is still in progress.

Acknowledgment. The results described in these lectures are taken from joint work with Anton Bovier, Alex Gaudillière, Dima Ioffe, Francesca Nardi, Enzo Olivieri, Elisabetta Scoppola and Cristian Spitoni. Part of this work was supported by the Dutch-German Bilateral Research Group "Random Spatial Models from Physics and Biology", which is jointly funded by DFG and NWO. These lectures were also presented at the "School on Information and Randomness", Center for Mathematical Modeling, University of Chile at Santiago, 11–15 December 2006.

2 Lecture 1: Finite Systems at Low Temperature, Definitions

In Lecture 1, we define two models: (I) Ising spins subject to Glauber dynamics; (II) lattice gas particles subject to Kawasaki dynamics. We fix the metastable regimes of interest and introduce the notions of communication height and communication level set between metastable states. In Lecture 2, we formulate two theorems for these two models that quantify their metastable behavior.

2.1 Glauber Dynamics and Kawasaki Dynamics

Let $\Lambda \subset \mathbb{Z}^2$ be a large finite box. We consider two types of configurations:

(I) *Ising spins*: $\eta = \{\eta(x)\colon\ x \in \Lambda\} \in \mathcal{X} = \{-1, +1\}^\Lambda$;
-1 = down-spin, $+1$ = up-spin (see Fig. 2).
(II) *Lattice gas*: $\eta = \{\eta(x)\colon\ x \in \Lambda\} \in \mathcal{X} = \{0, 1\}^\Lambda$;
0 = vacant, 1 = occupied (see Fig. 2).

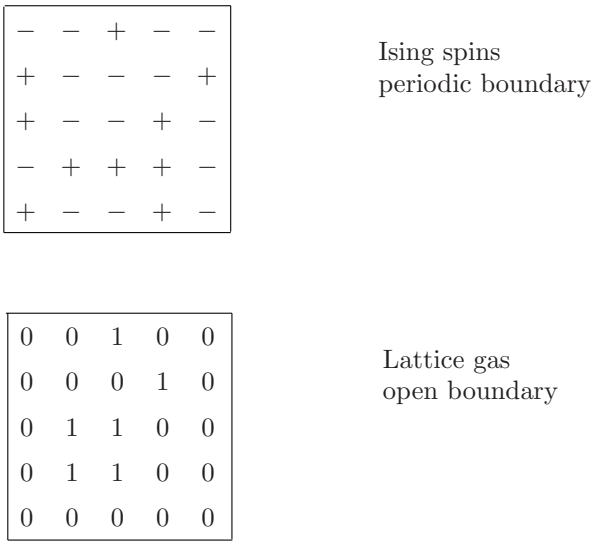


Fig. 2. Configurations of Ising spins and lattice gas.

On the configuration space $\mathcal{X}$, we consider the following Hamiltonians assigning an energy to each configuration:

$$\begin{aligned} &\text{(I)}: \ H(\eta) = -\tfrac{J}{2} \sum_{\substack{x,y \in \Lambda \\ x \sim y}} \eta(x)\eta(y) - \tfrac{h}{2} \sum_{x \in \Lambda} \eta(x), \\ &\text{(II)}: H(\eta) = -U \sum_{\substack{x,y \in \mathrm{int}(\Lambda) \\ x \sim y}} \eta(x)\eta(y) + \Delta \sum_{x \in \Lambda} \eta(x), \end{aligned} \tag{2.1}$$

where $\mathrm{int}(\Lambda) = \Lambda \backslash \partial\Lambda$ and $x \sim y$ means that x and y are neighboring sites. In (I) we pick *periodic* boundary conditions, in (II) we pick *open* boundary conditions (see (2.5)–(2.6) below). The parameters are:

(I) $J > 0$ the *ferromagnetic pair potential* and $h > 0$ the *magnetic field*;
(II) $U > 0$ the *binding energy* and $\Delta > 0$ the *activation energy*.

Definition 2.1. *The Metropolis dynamics at inverse temperature* $\beta \in (0,\infty)$ *is the continuous-time Markov process* $X = (X(t))_{t\geq 0}$ *on* $\mathcal{X}$ *with transition rates*

$$c(\eta,\eta') = \exp\left\{-\beta[H(\eta') - H(\eta)]_+\right\}, \qquad \eta,\eta' \in \mathcal{X}, \tag{2.2}$$

(where $[\cdot]_+$ *denotes the positive part) and allowed transitions*

$$\begin{aligned} &\text{(I)}: \ \eta' = \eta^x, \quad x \in \Lambda, \\ &\text{(II)}: \eta' = \eta^{x,y},\ x,y \in \Lambda,\ x \sim y, \end{aligned} \tag{2.3}$$

where

$$\eta^x(y) = \begin{cases} \eta(y), & y \neq x, \\ -\eta(x), & y = x, \end{cases} \qquad \eta^{x,y}(z) = \begin{cases} \eta(z), & z \neq x,y, \\ \eta(x), & z = y, \\ \eta(y), & z = x. \end{cases} \tag{2.4}$$

In words, for Ising spins the dynamics consists of spin-flips at single sites, called *Glauber dynamics*, while for the lattice gas it consists of exchange of occupation numbers between neighboring sites, called *Kawasaki dynamics*.

In the second dynamics, we also allow particles to *enter* and *exit* at $\partial\Lambda$. To that end, we also allow transitions

$$\text{(II)}: \quad \eta' = \eta^{*,x}, \quad x \in \partial\Lambda, \tag{2.5}$$

where

$$\eta^{*,x}(y) = \begin{cases} \eta(y), & y \neq x, \\ 1-\eta(x), & y = x. \end{cases} \tag{2.6}$$

View this as mimicking the presence of an *infinite gas reservoir* in $\mathbb{Z}^2\backslash\Lambda$ with density $e^{-\beta\Delta}$, which inserts particles at the sites of $\partial\Lambda$ at rate $e^{-\beta\Delta}$ and removes particles from the sites of $\partial\Lambda$ at rate 1.

A key observation is the following. The Metropolis dynamics has the *Gibbs measure*

$$\mu(\eta) = \frac{1}{Z}\, e^{-\beta H(\eta)}, \qquad \eta \in \mathcal{X}, \tag{2.7}$$

with Z the normalizing partition sum, as its *reversible equilibrium*, i.e.,

$$\mu(\eta)c(\eta,\eta') = \mu(\eta')c(\eta',\eta), \qquad \forall\, \eta,\eta' \in \mathcal{X}. \tag{2.8}$$

Note that the two Hamiltonians in (2.1) can be transformed into each other via the transformation

$$\eta(x) \leftrightarrow \frac{1}{2}[1+\eta(x)], \quad h \leftrightarrow 2U-\Delta, \quad J \leftrightarrow \frac{1}{2}U \tag{2.9}$$

(modulo constant terms and boundary terms). However, the allowed transitions for the two dynamics cannot be transformed into each other. Indeed, the first dynamics is *non-conservative*, the second dynamics is *conservative* (except at the boundary $\partial\Lambda$).

In what follows, we write $\mathbb{P}_\eta$ to denote the law of X given $X(0)=\eta$. For $\mathcal{A} \subset \mathcal{X}$, we write

$$\tau_{\mathcal{A}} = \inf\{t \geq 0 \colon\ X(t) \in \mathcal{A},\ X(t-) \notin \mathcal{A}\} \tag{2.10}$$

to denote the first entrance time of $\mathcal{A}$ by X.

2.2 Metastable Regimes

We will study the two dynamics in the *low temperature* limit $\beta \to \infty$, in their so-called *metastable regimes*:

$$\text{(I)}\colon\ 0<h<2J, \qquad \text{(II)}\colon\ U<\Delta<2U. \tag{2.11}$$

The dynamics will start in the configurations

$$\text{(I)}\colon\ \boxminus = \text{all spins down}, \qquad \text{(II)}\colon\ \square = \text{all sites vacant}, \tag{2.12}$$

and we will be interested in how the dynamics tunnels to the configurations

$$\text{(I)}\colon\ \boxplus = \text{all spins up}, \qquad \text{(II)}\colon\ \blacksquare = \text{all sites occupied}. \tag{2.13}$$

To understand the restrictions in (2.11), let us consider the energy of an $\ell \times \ell$ droplet inside Λ, i.e.,

$$\text{(I)}\colon\ E(\ell) = H(\eta_{\ell\times\ell}) - H(\boxminus), \qquad \text{(II)}\colon\ E(\ell) = H(\eta_{\ell\times\ell}) - H(\square). \tag{2.14}$$

(Note that $H(\square)=0$, while $H(\boxminus)<0$ when Λ is large enough depending on J and h.) An easy computation gives

$$\text{(I)}\colon\ E(\ell) = J[4\ell] - h\ell^2, \qquad \text{(II)}\colon\ E(\ell) = -U[2\ell(\ell-1)] + \Delta\ell^2. \tag{2.15}$$

In both cases, $\ell \mapsto E(\ell)$ is a downward parabola that goes through a maximum at $\ell = \frac{2J}{h}$, respectively, $\ell = \frac{U}{2U-\Delta}$. Hence, if both these ratios are non-integer, then the *critical droplets* (i.e., the droplets with

maximal energy on the parabola) are somewhere between a square of size $\ell_c - 1$ and a square of size ℓ_c, where

$$\text{(I)}\colon\ \ell_c = \left\lceil \frac{2J}{h} \right\rceil, \qquad \text{(II)}\colon\ \ell_c = \left\lceil \frac{U}{2U - \Delta} \right\rceil, \tag{2.16}$$

are the *critical droplet sizes*. The regimes in (2.11) correspond to $\ell_c \in (1, \infty)$.

In configuration space, we have the following qualitative picture:

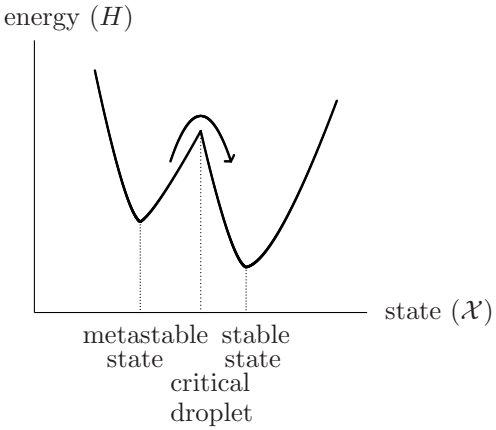


Fig. 3. Qualitative energy landscape.

The metastable regimes in (2.11) correspond to the situation where ⊟ and □ are *local minima* of the energy (lying at the bottom of a larger valley), ⊞ and ■ are *global minima*, and for the dynamics to move from one to the other it has to "go over an energetic hill". In physics language, we say that ⊟ and □ are *metastable states*, ⊞ and ■ are *stable states*, and the top of the hill separating them are *critical droplets*. We will address the following two questions (see Fig. 3):

(A) What are the *critical droplets* for the transitions ⊟ → ⊞ and □ → ■?

(B) How large are the *crossover times* $\tau_{\boxplus}$ and $\tau_{\blacksquare}$ starting from ⊟ and □, respectively?

2.3 Communication Height and Level Set

Write $\omega\colon\ \eta \to \eta'$ to denote a path of allowed transitions from η to η'.

Definition 2.2. *The communication height between* $\boxminus$ *and* $\boxplus$ *is defined as*

$$\text{(I)}\colon\ \Gamma = \Gamma(\boxminus, \boxplus) = \min_{\omega\colon\ \boxminus \to \boxplus} \max_{\xi \in \omega} [H(\xi) - H(\boxminus)]. \tag{2.17}$$

The corresponding communication level set is

$$\text{(I)}\colon\ \mathcal{S} = \mathcal{S}(\boxminus, \boxplus) = \Big\{\zeta \in \mathcal{X}\colon\ \exists\, \omega\colon\ \boxminus \to \boxplus \text{ with } \omega \ni \zeta \text{ such that } \max_{\xi \in \omega} [H(\xi) - H(\boxminus)] = H(\zeta) - H(\boxminus) = \Gamma\Big\}. \tag{2.18}$$

Similar definitions apply for $\square, \blacksquare$ *(with* $H(\square) = 0$*).*

In words, Γ is the minimal amount the energy has to increase in a path that achieves the crossover, called the *activation energy*, while $\mathcal{S}$ is the set of all *saddle point* configurations in the path (recall Fig. 3).

Our *intuitive guess* for the answer to question (A) is that the critical droplets are the configurations in $\mathcal{S}$, and for the answer to question (B) that

$$\tau_{\boxplus},\ \tau_{\blacksquare} \approx e^{\beta\Gamma} \quad \text{as } \beta \to \infty. \tag{2.19}$$

We will show in Lecture 2 that (2.19) is correct, obtaining in fact sharp estimates on $\mathbb{E}_{\boxminus}(\tau_{\boxplus})$ and $\mathbb{E}_{\square}(\tau_{\blacksquare})$, but that the critical droplets actually form a smaller set of configurations than $\mathcal{S}$, with an interesting geometry. We will see that models (I) and (II) show interesting similarities and differences.

3 Lecture 2: Finite Systems at Low Temperature, Theorems

In this lecture, we formulate two theorems that quantify the metastable behavior of models (I) and (II) in the regimes (2.11) by providing detailed answers to questions (A) and (B).

3.1 Glauber Dynamics

Theorem 3.1. (Neves and Schonmann [22], Bovier and Manzo [10])
(a) There exists a set of configurations $\mathcal{C}^* \subsetneq \mathcal{S}$ *such that*

$$\lim_{\beta\to\infty} \mathbb{P}_{\boxminus}(\tau_{\mathcal{C}^*} < \tau_{\boxplus} \mid \tau_{\boxplus} < \tau_{\boxminus}) = 1, \tag{3.1}$$

while this fails for any smaller set.
(b) The configurations in $\mathcal{C}^*$ *are those where the up-spins form an* $\ell_c \times (\ell_c - 1)$ *quasi-square, with a single protuberance on one of the sides of length* ℓ_c*.*

(c) The entrance distribution on $\mathcal{C}^*$ is asymptotically uniform:

$$\lim_{\beta\to\infty} \mathbb{P}_{\boxminus}\left(X(\tau_{\mathcal{C}^*}) = \eta \mid \tau_{\mathcal{C}^*} < \tau_{\boxplus}\right) = |\mathcal{C}^*|^{-1} \quad \forall\, \eta \in \mathcal{C}^*. \tag{3.2}$$

(d) There exists a constant $0 < K = K(\Lambda, \ell_c) < \infty$ such that

$$\lim_{\beta\to\infty} e^{-\beta\Gamma} \mathbb{E}_{\boxminus}(\tau_{\boxplus}) = K \tag{3.3}$$

with

$$\Gamma = H(\mathcal{C}^*) = J[4\ell_c] - h[\ell_c(\ell_c - 1) + 1], \tag{3.4}$$

and

$$\lim_{\beta\to\infty} \mathbb{P}_{\boxminus}(\tau_{\boxplus} > t\mathbb{E}_{\boxminus}(\tau_{\boxplus})) = e^{-t} \qquad \forall\, t \geq 0. \tag{3.5}$$

(e) For all Λ,

$$K(\Lambda, \ell_c) = \frac{3}{4(2\ell_c - 1)} \frac{1}{|\Lambda|}. \tag{3.6}$$

Parts (a)–(b), together with the crude estimate $\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{E}_{\boxminus}(\tau_{\boxplus}) = \Gamma$, were proved in [22]. Parts (c)–(e) were proved in [10].

Theorem 3.1(a) says that the configurations in $\mathcal{C}^*$ are the *critical droplets* that represent the *gate* for the crossover. According to Theorem 3.1(c), the entrance distribution of this gate is uniform.

Theorem 3.1(b) is explained as follows (see Fig. 5). Since the dynamics flips one spin at a time, on its way from $\boxminus$ to $\boxplus$ it must pass through a configuration that has $\ell_c(\ell_c - 1)$ up-spins. Among the configurations with precisely this number of up-spins, those where the up-spins form an $\ell_c \times (\ell_c - 1)$ quasi-square (of any location and orientation) have the smallest energy (due to a discrete isoperimetric inequality; see e.g. Alonso and Cerf [1]). Continuing on its way from $\boxminus$ to $\boxplus$, the dynamics must flip one more spin upwards. The configurations with smallest energy are those where this spin is attached to one of the sides of the quasi-square, forming a protuberance (see Fig. 4). Next, if this protuberance sits on one of the sides of length ℓ_c, then the dynamics can proceed downwards in energy by successively flipping up the spins next to the protuberance, to end up in an $\ell_c \times \ell_c$ square. This square is "over the hill" (see Fig. 3), because both its side lengths are supercritical (recall (2.16)). On the other hand, if the protuberance sits on one of the sides of length $\ell_c - 1$, then the dynamics can proceed downwards in energy to form an $(\ell_c - 1) \times (\ell_c + 1)$ rectangle, but this rectangle is "not over the hill", because one of its side lengths is subcritical.

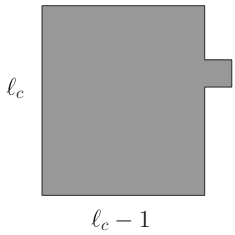


Fig. 4. A critical droplet for model (I): The up-spins lie inside the shaded area, the down-spins outside.

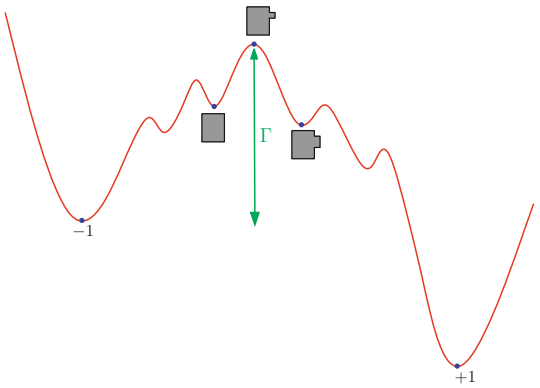


Fig. 5. A nucleation path.

Here is an explanation of Theorem 3.1(d–e). The exponential law comes from the fact that the crossover only occurs after many unsuccessful attempts to create a critical droplet and "go over the hill". The average time needed to enter $\mathcal{C}^*$ is

$$\frac{1}{|\mathcal{C}^*|}\, e^{\beta\Gamma}\,[1+o(1)] \qquad \text{as } \beta\to\infty. \tag{3.7}$$

Let $\pi(\ell_c)$ denote the average probability with respect to the uniform entrance distribution that the critical droplet is exited in the direction of $\boxplus$ rather than $\boxminus$. Then the average number of attempts to go over the hill after reaching the top is

$$\frac{1}{\pi(\ell_c)}\,[1+o(1)] \qquad \text{as } \beta\to\infty. \tag{3.8}$$

The product of (3.7) and (3.8) is the average crossover time, and so

$$K = \frac{1}{|\mathcal{C}^*|\pi(\ell_c)}. \tag{3.9}$$

Now,

$$|\mathcal{C}^*| = |\Lambda|\, 4\ell_c, \tag{3.10}$$

because the droplet can be centered anywhere in Λ, has 2 possible orientations, and the protuberance can sit in $2\ell_c$ places. Moreover,

$$\pi(\ell_c) = \frac{1}{\ell_c}\left(2\,\frac{1}{2} + (\ell_c - 2)\,\frac{2}{3}\right). \tag{3.11}$$

Indeed, if the protuberance sits at one of the two extreme ends of a side of length ℓ_c, then the probability is $\frac{1}{2}$ that its *one* neighboring spin on the same side flips upwards before the protuberance flips downwards. On the other hand, when the protuberance sits at one of the $\ell_c - 2$ other locations on this side, then it has *two* neighboring spins on the same side and so the probability for one of them to flip upwards before the protuberance flips downwards is $\frac{2}{3}$. Combining (3.9–3.11), we get (3.6).

In Theorem 3.1(a), an example of a configuration in $\mathcal{S}\backslash\mathcal{C}^*$ is obtained by picking any configuration in $\mathcal{C}^*$, flipping up any spin next to the protuberance (at gain h) and afterwards flipping down any spin at a corner of the quasi-square (at cost h). For the dynamics, this configuration is a *dead-end*. Indeed, the last flip must be reversed before the dynamics can initiate the motion downhill to the $\ell_c \times \ell_c$ square.

3.2 Kawasaki Dynamics

Theorem 3.2. (den Hollander, Olivieri and Scoppola [19], Bovier, den Hollander and Nardi [8])
(a) There exists a set of configurations $\mathcal{C}^* \subsetneq \mathcal{S}$ such that

$$\lim_{\beta\to\infty} \mathbb{P}_{\square}(\tau_{\mathcal{C}^*} < \tau_{\blacksquare} \mid \tau_{\blacksquare} < \tau_{\square}) = 1, \tag{3.12}$$

while this fails for any smaller set.
(b) The configurations in $\mathcal{C}^*$ are those where the particles either form an $(\ell_c - 2) \times (\ell_c - 2)$ square, with four bars attached to the four sides of total length $3\ell_c - 3$ and 1 free particle, or form an $(\ell_c - 1) \times (\ell_c - 3)$ rectangle, with four bars attached to the four sides of total length $3\ell_c - 2$ and 1 free particle.
(c) The entrance distribution on $\mathcal{C}^*$ is asymptotically uniform:

$$\lim_{\beta\to\infty} \mathbb{P}_{\square}\left(X(\tau_{\mathcal{C}^*}) = \eta \mid \tau_{\mathcal{C}^*} < \tau_{\square}\right) = |\mathcal{C}^*|^{-1} \quad \forall\, \eta \in \mathcal{C}^*. \tag{3.13}$$

(d) There exists a constant $0 < K = K(\Lambda, \ell_c) < \infty$ such that

$$\lim_{\beta\to\infty} e^{-\beta\Gamma}\mathbb{E}_{\square}(\tau_{\blacksquare}) = K \tag{3.14}$$

with

$$\Gamma = H(\mathcal{C}^*) = -U(2\ell_c^2 - 4\ell_c + 2) + \Delta(\ell_c^2 - \ell_c + 2), \tag{3.15}$$

and

$$\lim_{\beta\to\infty} \mathbb{P}_{\square}(\tau_{\blacksquare} > t\mathbb{E}_{\square}(\tau_{\blacksquare})) = e^{-t} \qquad \forall\, t \geq 0. \tag{3.16}$$

(e) As $\Lambda \to \mathbb{Z}^2$,

$$K(\Lambda, \ell_c) \sim \frac{3}{4\pi\ell_c^2(\ell_c^2 - 1)} \frac{\log|\Lambda|}{|\Lambda|}. \tag{3.17}$$

Part (a), together with a partial description of $\mathcal{C}^*$ and the crude estimate $\lim_{\beta\to\infty}(1/\beta)\log \mathbb{E}_{\boxminus}(\tau_{\boxplus}) = \Gamma$, were proved in [19]. Parts (b)–(e) were proved in [8].

Comparing Theorem 3.2(b) with Theorem 3.1(b), we see that the critical droplet for Kawasaki is more complicated than for Glauber. Once the dynamics has created a *protocritical droplet* (= quasi-square plus protuberance without free particle), it must wait for the next particle to arrive from the boundary (which is the free particle in Fig. 6). This takes a time of order $e^{\beta\Delta}$. Because $\Delta > U$, this time is much larger than $e^{\beta U}$, the time for the dynamics to make moves that cost U. Therefore the droplet will "explore" all shapes that can be reached from its protocritical shape via a U-path, i.e., a path between two configurations with the same energy that never goes more than U above this energy. For instance, the protuberance may detach itself from the side of length ℓ_c and reattach itself to the side of length $\ell_c - 1$. But it is also possible for particles to *slide along the boundary of the droplet*, in a train-like motion around corners (see Fig. 7), so as to modify the four bars in the annulus of the droplet.

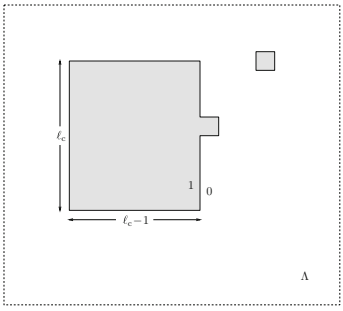


Fig. 6. A critical droplet for model (II): The occupied sites lie inside the shaded areas, the vacant sites outside.

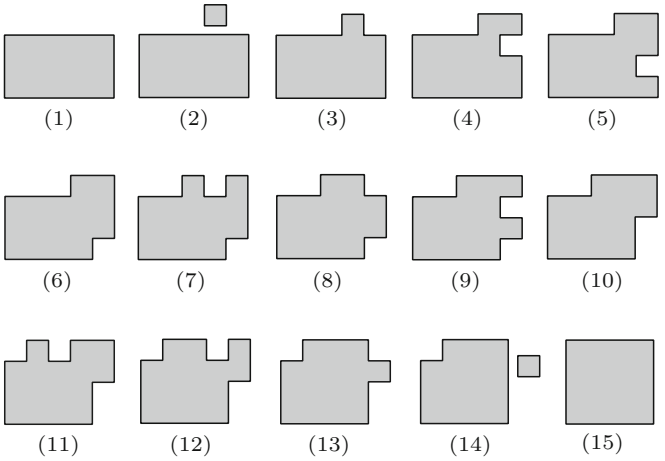


Fig. 7. Motion along the border of the droplet. Configurations (3–13) form a U-path.

Theorem 3.1(d–e) is explained as follows. Write $\mathcal{C}$ to denote the set of protocritical droplets. The average time needed to enter $\mathcal{C}^*$ is

$$\frac{1}{|\mathcal{C}|\,|\partial\Lambda|}\,e^{\beta\Gamma}\,[1+o(1)] \quad \text{as } \beta\to\infty. \tag{3.18}$$

Let $\pi(\Lambda,\ell_c)$ denote the average probability with respect to the uniform entrance distribution that the critical droplet is exited in the direction of ■ rather than □. Then the average number of attempts to go over the hill after reaching the top is

$$\frac{1}{\pi(\Lambda,\ell_c)}\,[1+o(1)] \quad \text{as } \beta\to\infty. \tag{3.19}$$

The product of (3.7) and (3.8) is the average crossover time, and so

$$K=\frac{1}{|\mathcal{C}|\,|\partial\Lambda|\,\pi(\Lambda,\ell_c)}. \tag{3.20}$$

Now,

$$|\mathcal{C}|\sim|\Lambda|\,\frac{1}{3}\ell_c^2(\ell_c^2-1) \quad \text{as } \Lambda\to\mathbb{Z}^2, \tag{3.21}$$

where the first factor comes from centering the droplet anywhere in Λ not touching $\partial\Lambda$, while the second factor comes from a combinatorial calculation counting the number of sizes and locations of the four bars in the annulus. Moreover,

$$|\partial\Lambda|\,\pi(\ell_c,\Lambda) \sim \frac{4\pi}{\log|\Lambda|} \quad \text{as } \Lambda \to \mathbb{Z}^2. \tag{3.22}$$

Indeed, the right-hand side is the probability that a particle detaching itself from the critical droplet reaches $\partial\Lambda$ and exits Λ before reattaching itself. This probability is asymptotically independent of the shape and the location of the critical droplet, due to the fact that the free particle moves like a two-dimensional simple random walk (which is recurrent on $\mathbb{Z}^2$). By reversibility, the reverse motion has the same probability, which is the left-hand side. Combining (3.20–3.22), we get (3.17).

In Theorem 3.1(a), an example of a configuration in $\mathcal{S}\backslash\mathcal{C}^*$ is an $\ell_c \times (\ell_c - 1)$ quasi-square plus a dimer at distance 1. For the dynamics, this configuration is a *dead-end*. Indeed, one particle of the dimer must jump back to the droplet and create a protuberance (at cost 0), and the remaining free particle must attach itself next to this protuberance (at gain U) to initiate the motion downhill to the $\ell_c \times \ell_c$ square.

3.3 Potential-Theoretic Approach

We give a sketch of the techniques that are used to obtain the fine asymptotics of the average crossover time in Theorems 3.1(d–e) and 3.2(d–e). A key role is played by the notion of capacity between two sets of configurations, in particular, between the metastable state and the stable state. We refer to Section 3–5 in Bovier [4] for the general background of this notion within the context of metastability and for other applications.

Define

$$\mathcal{E}(h) = \frac{1}{2} \sum_{\eta,\eta' \in \mathcal{X}} \mu(\eta) c(\eta,\eta') [h(\eta) - h(\eta')]^2, \qquad h\colon\ \mathcal{X} \to [0,1]. \tag{3.23}$$

This is the *Dirichlet form* associated with the dynamics, whose argument is a *potential function* on the configuration space $\mathcal{X}$. Given two non-empty disjoint sets $\mathcal{A}, \mathcal{B} \subset \mathcal{X}$, the *capacity* of the pair $\mathcal{A}, \mathcal{B}$ is defined as

$$\mathrm{CAP}(\mathcal{A},\mathcal{B}) = \min_{\substack{h\colon\ \mathcal{X}\to[0,1] \\ h|_{\mathcal{A}}\equiv 1, h|_{\mathcal{B}}\equiv 0}} \mathcal{E}(h), \tag{3.24}$$

where the minimum runs over all potential functions whose restriction to $\mathcal{A}$ and $\mathcal{B}$ equals 1 and 0, respectively. If we think of an electric network with nodes labelled by $\mathcal{X}$ and with conductivities $\mu(\eta)c(\eta,\eta')$ between nodes $\eta, \eta' \in \mathcal{X}$, then $\mathcal{E}(h)$ is the energy produced by an electric current flowing through this network when the potential on the nodes

is given by h. The capacity is the minimal energy when the nodes of $\mathcal{A}$ are kept at potential 1 and the nodes of $\mathcal{B}$ are kept at potential 0 ("Thompson's principle"). The minimum in (3.24) is unique, and the minimizer h^* has the interpretation

$$h^*(\eta) = \mathbb{P}_\eta(\tau_{\mathcal{A}} < \tau_{\mathcal{B}}) \qquad \text{for } \eta \notin A \cup B. \tag{3.25}$$

What is important about (3.24) is that upper bounds can be obtained by inserting test functions for h, while lower bounds can be obtained by removing transitions from $\mathcal{X} \times \mathcal{X}$ ("Rayleigh's short-cut rule"). This gives great flexibility in the calculations.

We henceforth focus on model (II), but the claims made below apply equally well to model (I). A key ingredient is the following fact, implying that $\{\square, \blacksquare\}$ is a *metastable pair* for low temperature.

Proposition 3.3. (den Hollander, Nardi, Olivieri and Scoppola [18])
For all $\eta \in \mathcal{X} \backslash \{\square, \blacksquare\}$,

$$\Gamma(\eta, \{\square, \blacksquare\}) < \Gamma, \tag{3.26}$$

where

$$\Gamma(\mathcal{A}, \mathcal{B}) = \min_{\eta \in \mathcal{A}, \eta' \in \mathcal{B}} \min_{\omega:\ \eta \to \eta'} \max_{\xi \in \omega} [H(\xi) - H(\eta)] \tag{3.27}$$

is the communication height between $\mathcal{A}, \mathcal{B} \subset \mathcal{X}$, $\mathcal{A} \cap \mathcal{B} = \emptyset$, $\mathcal{A}, \mathcal{B} \neq \emptyset$.

Proposition 3.3 implies that no matter where the dynamics starts, it reaches the set $\{\square, \blacksquare\}$ faster than it manages to achieve the crossover from $\square$ to $\blacksquare$. In words, there are "no deep pits" in the energy landscape that trap the dynamics for a time comparable to the crossover time.

The key to the fine estimate in Theorem 3.2(d–e) is the following fact, relating the average crossover time to the capacity and relying crucially on Proposition 3.3.

Proposition 3.4. (Bovier, den Hollander and Nardi [8])
$\mathbb{E}_\square(\tau_\blacksquare) = [1 + o(1)]/[Z\mathrm{CAP}(\square, \blacksquare)]$ *as* $\beta \to \infty$.

Thus, to estimate the average crossover time from $\square$ to $\blacksquare$, it suffices to estimate the capacity of the pair $\square, \blacksquare$. This proceeds in several steps.

(1) A crude *a priori estimate* yields that for every pair $\mathcal{A}, \mathcal{B}$ there exist constants $0 < C_1 < C_2 < \infty$ (depending on $\mathcal{A}, \mathcal{B}$ but not on β) such that

$$C_1 \leq e^{\beta\Gamma(\mathcal{A},\mathcal{B})} Z\mathrm{CAP}(\mathcal{A}, \mathcal{B}) \leq C_2. \tag{3.28}$$

The lower bound is obtained by picking a minimax path ω in (3.27) and from (3.23) remove all transitions $\eta \to \eta'$ that are not in ω. The

upper bound is obtained by picking a test function h in (3.24) that is $\equiv 1$ on the Γ-valley around $\mathcal{A}$ and $\equiv 0$ on the Γ-valley around $\mathcal{B}$. (The Γ-valley around a set of configurations $\mathcal{S}$ is the set of configurations $\mathcal{S}' \supsetneq \mathcal{S}$ whose communication height with $\mathcal{S}$ is $< \Gamma$.)

(2) With the help of (3.28), it is possible to obtain sharp bounds on the minimizer h^* of (3.24) given in (3.25) via so-called renewal-type estimates. These estimates show that h^* is exponentially close (in β) to 1 on the Γ-valley around $\square$ and exponentially close (in β) to 0 on the Γ-valley around $\blacksquare$. Since the configurations with energy $> \Gamma$ are negligible, because of the Gibbs factor in (3.23) coming form (2.7), it follows that the sharp asymptotics of $\text{CAP}(\square, \blacksquare) = \mathcal{E}(h^*)$ is determined by the values of h^* on $\mathcal{S} = \mathcal{S}(\square, \blacksquare)$ and on $\partial^{ext}\mathcal{S}$, the exterior boundary of $\mathcal{S}$.

(3) Due to the above, the variational problem in (3.24) on the full configuration space $\mathcal{X}$ reduces to a variational problem restricted to $\mathcal{S} \cup \partial^{ext}\mathcal{S}$. This *reduced variational problem* has a much simpler structure, and can be understood in terms of the geometry of the configurations that are critical droplets or are close to critical droplets.

(4) For Kawasaki, the reduced variational problem involves the creation of a free particle when the droplet is protocritical, the motion of this free particle towards the droplet, and the attachment itself. Since this is a problem about simple random walk travelling between $\partial\Lambda$ and a protocritical droplet somewhere inside Λ, the reduced capacity can be sharply estimated.

For Kawasaki, $\mathcal{S} \cup \partial^{ext}\mathcal{S}$ contains plateaus, wells and dead-ends, and hence a closed form computation of $K = K(\ell_c, \Lambda)$ is not feasible. Fortunately, for large Λ the details of the geometry of $\mathcal{S} \cup \partial^{ext}\mathcal{S}$ turn out to be only partly relevant, and the asymptotics of K can be identified, resulting in (3.17). For Glauber, the reduced variational problem turns out to be zero-dimensional, and $K = K(\ell_c, \Lambda)$ can be computed in closed form, resulting in (3.6).

Remark: Most of the results described above can be extended to *other types of dynamics*, such as Glauber dynamics for Ising spins with an anisotropic interaction or in a staggered magnetic field, or Ising spins subject to a parallel dynamics given by a probabilistic cellular automaton (see den Hollander [17] for references). Similarly, most of the results can be extended to three dimensions, despite the more complex geometry of critical droplets (see Ben Arous and Cerf [2], den Hollander, Nardi, Olivieri and Scoppola [18] for the necessary background).

4 Lecture 3: Large Systems at Low Temperature and Moderate Systems at Positive Temperature, Conjectures

In this lecture we move away from finite systems and investigate what happens in growing volumes, both at low and at positive temperature. Most of what is described below consists of target theorems and work in progress.

Glauber dynamics for large systems at low temperature was studied in Dehghanpour and Schonmann [12], [13], Schonmann and Shlosman [30], and Manzo and Olivieri [21], using the *pathwise approach*. Current work focusses on trying to improve their results using the *potential-theoretic approach*, and on extending the analysis to Kawasaki dynamics.

4.1 Large Systems at Low Temperature

Glauber Dynamics

Let $\Lambda = \Lambda_\beta$ depend on β such that

$$|\Lambda_\beta| = e^{\Theta\beta}, \qquad \Theta \in [0, \infty). \tag{4.1}$$

Let

– $\mathcal{R} \subset \mathcal{X}$ denote those configurations where the *circumscribed rectangles* of all clusters of up-spins in Λ_β are contained in *non-interacting protocritical quasi-squares* (see Fig. 8).

The initial configuration $X(0)$ of the dynamics is drawn according to the *conditional* Gibbs measure

$$\mu_{\mathcal{R}}(\eta) = \frac{\mu(\eta)1_{\mathcal{R}}(\eta)}{\mu(\mathcal{R})}, \qquad \eta \in \mathcal{X}_\beta, \tag{4.2}$$

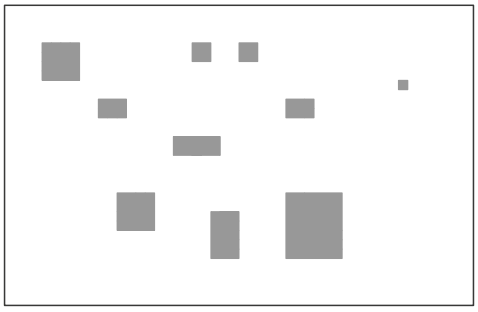


Fig. 8. A configuration in $\mathcal{R}$.

where μ is defined in (2.7) and $\mu(\mathcal{R}) = \sum_{\eta\in\mathcal{R}} \mu(\eta)$. Our goal will be to estimate the first time a critical droplet appears *anywhere* in Λ_β, i.e., the first exit time of $\mathcal{R}$.

As before, we will be interested in the metastable regime

$$h \in (0, 2J), \quad \beta \to \infty. \tag{4.3}$$

Write $\mathbb{P}_\eta$ to denote the law of the dynamics $X = (X(t))_{t\geq 0}$ starting from $X(0) = \eta$, and put $\mathbb{P}_{\mu_\mathcal{R}} = \sum_{\eta\in\mathcal{R}} \mu_\mathcal{R}(\eta)\mathbb{P}_\eta$. Let

$$\tau_{\mathcal{R}^c} = \min\{t \geq 0\colon\ X(t) \notin \mathcal{R}\} \tag{4.4}$$

denote the first time the dynamics exits $\mathcal{R}$. Write $\asymp$ for asymptotic equality modulo constants.

Conjecture 4.1. (Bovier, den Hollander and Spitoni [9])
If

$$\Theta \in [0, \Gamma - \Xi) \quad \textit{with} \quad \Xi = h(\ell_c - 2), \tag{4.5}$$

then

$$\mathbb{E}_{\mu_\mathcal{R}}(\tau_{\mathcal{R}^c}) \asymp \frac{1}{|\Lambda_\beta|}\, e^{\beta\Gamma} \qquad \textit{as } \beta \to \infty. \tag{4.6}$$

The idea behind Conjecture 4.1 is simple. The dynamics grows and shrinks droplets essentially independently in different local boxes. Consequently, a critical droplet appears randomly in one of the local boxes, after a time that is the local crossover time divided by the number of local boxes in $|\Lambda_\beta|$. This is the regime of *homogeneous nucleation*.

Γ is the *local energy of the critical droplet*, which plays the role of the local activation energy for the crossover. Ξ is the *local energy needed to evaporate the largest subcritical droplet*. The regime in (4.5) corresponds to the situation where any subcritical droplet has a tendency to evaporate in a time much smaller than the crossover time.

Kawasaki Dynamics

Keep (4.1). This time, let

– $\mathcal{R} \subset \mathcal{X}$ denote those configurations where all clusters of particles in Λ_β are either *strictly contained in* a protocritical quasi-square plus protuberance (recall Fig. 6), or *are equal to* a protocritical quasi-square plus protuberance with an *empty annulus* Λ^* of a size slightly smaller than the typical interparticle distance (see Fig. 9).

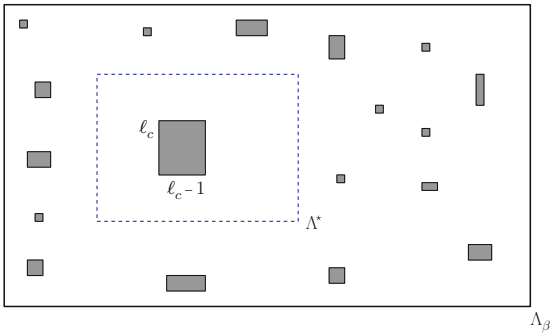


Fig. 9. A configuration in $\mathcal{R}$.

The initial configuration $X(0)$ of the dynamics is again drawn according to (4.2). We run the dynamics associated with the Hamiltonian in the second line of (2.1) *without the activity term*. Indeed, this term is no longer needed, because Λ_β *is so large that it takes over the role of the gas reservoir*. In fact, we will supply Λ_β with periodic boundary conditions, so that no particle enters or exits Λ_β at positive times. Our choice to start from the conditional equilibrium *with activity*, given by (4.2), is needed at time zero only, and is made for convenience. Thus, the particle density inside Λ_β is $e^{-\beta\Delta}$ at time zero and remains fixed in the course of time.

In order to have particles at all we must pick $\Theta > \Delta$. We will be interested in the regime

$$\Delta \in (U, 2U), \quad \Theta \in (\Delta, \infty), \quad \beta \to \infty. \tag{4.7}$$

Conjecture 4.2. (Gaudilliere, den Hollander, Nardi, Olivieri and Scoppola [16], Bovier, den Hollander and Spitoni [9])
Suppose that $\ell_c \geq 3$. *If*

$$\Theta \in (\Delta, \Gamma - \Xi) \quad \textit{with} \quad \Xi = 2U + (\ell_c - 3)(2U - \Delta), \tag{4.8}$$

then

$$\mathbb{E}_{\mu_{\mathcal{R}}}(\tau_{\mathcal{R}^c}) \asymp \beta \frac{1}{|\Lambda_\beta|} e^{\beta\Gamma} \qquad \textit{as } \beta \to \infty. \tag{4.9}$$

Because of the low particle density, as before, droplets grow and shrink more or less independently in different local boxes, causing homogeneous nucleation.

4.2 Moderate Systems at Positive Temperature

In *equilibrium* statistical physics, for a system that is at a *first-order* phase transition a macroscopically large droplet of one phase inside the other phase takes on the *Wulff shape*, i.e., the droplet minimizes its total surface tension subject to a total volume constraint. This observation, which is over a century old, has been put on a rigorous microscopic basis since only fifteen years or so. For the two-dimensional ferromagnetic nearest-neighbor Ising model at low temperature, Dobrushin, Kotecký and Shlosman [14] proved that a large droplet of the plus-phase inside the minus-phase has the Wulff shape. This result was subsequently extended up to the critical temperature, and its proof was simplified, by Pfister [26], Ioffe [20] and Pisztora [27].

The Wulff construction requires a careful *coarse-graining* analysis. The microscopic phase boundary is approximated on a mesoscopic scale by a polygon consisting of many segments, which decouple on the mesoscopic scale. Each segment contributes to the surface tension in a way that depends on its direction relative to the lattice axes. To handle the fluctuations of the boundary around the polygon, large deviation arguments are required. The polygon tends to a smooth curve in the macroscopic limit, and this curve enters into the *Wulff variational problem*, whose solution is the actual phase boundary.

To study Wulff droplets in the presence of a stochastic dynamics is part of *non-equilibrium* statistical physics and therefore is quite a different matter. The question of interest is whether *macroscopically large critical droplets for metastable transitions between two phases under a stochastic local dynamics assume the Wulff shape or not.*

In this lecture we allow the box Λ to grow but only moderately, in a way that depends not on β but on the parameters in the Hamiltonian.

Glauber Dynamics

We suppose that $\Lambda = \Lambda_h$ with

$$|\Lambda_h| = C\,\frac{1}{h}, \qquad 1 \ll C < \infty. \tag{4.10}$$

We assume that $\beta > \beta_c$, the critical inverse temperature at $h = 0$ for $\Lambda = \mathbb{Z}^2$. The system starts at $X(0) = \boxminus$, the Glauber dynamics is applied for small $h > 0$, and the limit $h \downarrow 0$ is taken. The dynamics eventually brings the system to equilibrium, close to the plus-phase at $h = 0$, but it needs a long time to do so. In the limit $h \downarrow 0$, the critical droplet becomes macroscopically large. The goal is to show that the

critical droplet scales to the *equilibrium Wulff shape* and appears after a time that scales like the exponential of the *Wulff free energy*. The size of the box is taken to scale in such a way that the critical droplet occupies a finite fraction of the box.

Theorem 4.3. (Schonmann and Shlosman [30])
For $\beta > \beta_c$ and C sufficiently large,

$$\lim_{h\downarrow 0} h \log \mathbb{E}_{\boxminus}(\tau_{m(J,\beta)}) = \frac{W(J,\beta)^2}{4m(J,\beta)}, \tag{4.11}$$

where $\tau_{m(J,\beta)}$ is the first time the total magnetization inside Λ_h equals $m(J,\beta)$, the spontaneous magnetization on $\mathbb{Z}^2$, and $W(J,\beta)$ is the total surface tension of the Wulff droplet of unit volume.

Note that the left-hand side of (4.11) refers to a non-equilibrium quantity, while the right-hand side only contains quantities from equilibrium. This is why the result in (4.11) is deep.

The idea behind (4.11) is that, in the macroscopic scaling limit, the critical droplet has a length ℓ that maximizes the free energy function

$$f(\ell) = -m(J,\beta)h\ell^2 + W(J,\beta)\ell. \tag{4.12}$$

(This is a macroscopic version of the parabola encountered in (2.15)!) The maximum is taken at $\ell_{max} = W(J,\beta)/2m(J,\beta)h$, giving free energy

$$f(\ell_{max}) = \frac{W(J,\beta)^2}{4m(J,\beta)h}. \tag{4.13}$$

This is the exponential of the time needed to create a droplet at a given location.

Schonmann and Shlosman [30] analyze the problem also on $\mathbb{Z}^2$ instead of on Λ_h subject to (4.10). They show that, in infinite volume, the critical droplet typically is not created close to the origin, but rather is *created far away and subsequently invades the origin by growing*. As a result, the exponential is *three times smaller*, because the critical droplet may occur anywhere in a space-time cone of this smaller size and invade the origin afterwards.

Kawasaki Dynamics

This time we suppose that $\Lambda = \Lambda_\Delta$ with

$$|\Lambda_\Delta| = C\,\frac{1}{2U-\Delta}, \qquad 1 \ll C < \infty. \tag{4.14}$$

We assume that $\beta > \beta_c$, the critical inverse temperature for the Hamiltonian without activity term for $\Lambda = \mathbb{Z}^2$. The system starts at $X(0) = \square$, the Kawasaki dynamics is applied for $\Delta < 2U$, and the limit $\Delta \uparrow 2U$ is taken. This is the limit of *weak supersaturation*, when the critical droplet becomes macroscopically large.

Conjecture 4.4. (Bovier, den Hollander and Ioffe [7])
For $\beta > \beta_c$ and C sufficiently large,

$$\lim_{\Delta \uparrow 2U} (2U - \Delta) \log \mathbb{E}_{\square}(\tau_{\rho(U,\beta)}) = \frac{W(U,\beta)^2}{2\rho(\beta,U) - 1}, \tag{4.15}$$

where $\tau_{\rho(U,\beta)}$ is the first time the particle density inside Λ_Δ equals $\rho(U,\beta)$, the density of the liquid phase on $\mathbb{Z}^2$, and $W(U,\beta)$ is the total surface tension of the Wulff droplet of unit volume.

The right-hand side of (4.15) is the same as that of (4.13), with J being replaced by $U/2$, because of the link between the Hamiltonians of models (I) and (II) in (2.1) (recall (2.9)). The reason is that, as already observed above, the right-hand side of (4.15) only contains quantities from *equilibrium*.

A proof of Conjecture 4.4 is currently being attempted, with the help of the potential-theoretic techniques mentioned in Section 3.3, for a simpler version of the model where the interaction is of *Kac-Dyson type*, i.e., quasi-mean-field. The hard part is that, for growing volumes at positive temperature, *both spatial and temporal entropy* need to be controlled. We need to understand the typical way in which the dynamics grows and shrinks large droplets, absorbing and emitting large numbers of particles with the surrounding gas phase in the box while keeping the droplet close to the Wulff shape. Droplets are expected to grow and shrink via "motion by curvature".

References

1. L. Alonso and R. Cerf, The three-dimensional polyominoes of minimal area, Electron. J. Combin. 3 (1996) Research Paper 27.
2. G. Ben Arous and R. Cerf, Metastability of the three-dimensional Ising model on a torus at very low temperature, Electron. J. Probab. 1 (1996) Research Paper 10.
3. A. Bovier, Metastability: a potential theoretic approach, Proceedings of ICM 2006, Vol. 3, EMS Publising House, 2006, pp. 499–518.

4. A. Bovier, Metastability. In this volume. Springer, Berlin, 2007.
5. A. Bovier, M. Eckhoff, V. Gayrard and M. Klein, Metastability in stochastic dynamics of disordered mean-field models, Probab. Theory Rel. Fields 119 (2001) 99–161.
6. A. Bovier, M. Eckhoff, V. Gayrard and M. Klein, Metastability and low lying spectra in reversible Markov chains, Commun. Math. Phys. 228 (2002) 219–255.
7. A. Bovier, F. den Hollander and D. Ioffe, work in progress.
8. A. Bovier, F. den Hollander and F.R. Nardi, Sharp asymptotics for Kawasaki dynamics on a finite box with open boundary, Probab. Theory Relat. Fields 135 (2006) 265–310.
9. A. Bovier, F. den Hollander and C. Spitoni, EURANDOM Report 2008-021.
10. A. Bovier and F. Manzo, Metastability in Glauber dynamics in the low-temperature limit: beyond exponential asymptotics, J. Stat. Phys. 107 (2002) 757–779.
11. M. Cassandro, A. Galves, E. Olivieri and M.E. Vares, Metastable behavior of stochastic dynamics: a pathwise approach, J. Stat. Phys. 35 (1984) 603–634.
12. P. Dehghanpour and R.H. Schonmann, Metropolis dynamics relaxation via nucleation and growth, Comm. Math. Phys. 188 (1997) 89–119.
13. P. Dehghanpour and R.H. Schonmann, A nucleation-and-growth model, Probab. Theory Rel. Fields 107 (1997) 123–135.
14. R.L. Dobrushin, R. Kotecký, S.B. Shlosman, *Wulff construction: a global shape from local interaction*, AMS translations series, Providence, RI, Am. Math. Soc. 1992.
15. M.I. Freidlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, Berlin, 1984.
16. A. Gaudillière, F. den Hollander, F.R. Nardi, E. Olivieri and E. Scoppola, work in progress.
17. F. den Hollander, Metastability under stochastic dynamics, Stoch. Proc. Appl. 114 (2004) 1–26.
18. F. den Hollander, F.R. Nardi, E. Olivieri and E. Scoppola, Droplet growth for three-dimensional Kawasaki dynamics, Probab. Theory Relat. Fields 125 (2003) 153–194.
19. F. den Hollander, E. Olivieri and E. Scoppola, Metastability and nucleation for conservative dynamics, J. Math. Phys. 41 (2000) 1424–1498.
20. D. Ioffe, Exact deviation bounds up to T_c for the Ising model in two dimensions, Probab. Theory Relat. Fields 102 (1995) 313–330.
21. F. Manzo and E. Olivieri, Dynamical Blume-Capel model: competing metastable states at infinite volume, J. Stat. Phys. 104 (2001) 1029–1090.
22. E.J. Neves and R.H. Schonmann, Critical droplets and metastability for Glauber dynamics at very low temperature, Comm. Math. Phys. 137 (1991) 209–230.

23. E. Olivieri and E. Scoppola, Metastability and typical exit paths in stochastic dynamics, Proceedings of ECM 1996, Progress in Mathematics 169, Birkhäuser, Basel, 1998, pp. 124–150.
24. E. Olivieri and M.E. Vares, *Large Deviations and Metastability*, Cambridge University Press, Cambridge, 2004.
25. O. Penrose and J.L. Lebowitz, Towards a rigorous molecular theory of metastability, in: *Fluctuation Phenomena*, 2nd. ed. (E.W. Montroll and J.L. Lebowitz, eds.), North-Holland, Amsterdam, 1987.
26. C.E. Pfister, Large deviations and phase separation in the two-dimensional Ising model, Helv. Phys. Acta 64 (1991) 953–1054.
27. A. Pisztora, Surface order large deviations for Ising, Potts and percolation models, Probab. Theory Relat. Fields 104 (1996) 427–466.
28. R.H. Schonmann, Theorems and conjectures on the droplet-driven relaxation of stochastic Ising models, in: *Probability and Phase Transitions* (G. Grimmett ed.), NATO ASI Series, Kluwer Academic Publishers, Dordrecht, 1994, pp. 265–301.
29. R.H. Schonmann, Metastability and the Ising model, Proceedings of ICM 1998, Doc. Math., Extra Vol. 3, pp. 173–181.
30. R.H. Schonmann, S.B. Shlosman, Wulff droplets and the metastable relaxation of kinetic Ising models, Comm. Math. Phys. 194 (1998) 389–462.
31. E. Scoppola, Metastability for Markov chains, in: *Probability and Phase Transitions* (G. Grimmett ed.), NATO ASI Series, Kluwer Academic Publishers, Dordrecht, 1994, pp. 303–322.
32. M.E. Vares, Large deviations and metastability, in: *Disordered Systems*, Traveaux en Cours 53, Hermann, Paris, 1996, pp. 1–62.

A Selection of Nonequilibrium Issues

Christian Maes[1], Karel Netočný[2], and Bidzina Shergelashvili[1*]

[1] Instituut voor Theoretische Fysica, K. U. Leuven, Belgium
Christian.Maes@fys.kuleuven.be
[2] Institute of Physics AS CR, Prague, Czech Republic
netocny@fzu.cz

Summary. We give a pedagogical introduction to a selection of recently discussed topics in nonequilibrium statistical mechanics, concentrating mostly on formal structures and on general principles. Part I contains an overview of the formalism of lattice gases that we use to explain various symmetries and inequalities generally valid for nonequilibrium systems, including the fluctuation symmetry, Jarzynski equality, and the direction of currents. That mostly concerns the time-antisymmetric part of dynamical fluctuation theory. We also briefly comment on recent attempts to combine that with the time-symmetric sector in a Langrangian or extended Onsager-Machlup approach. In Part II we concentrate on the macroscopic state and how entropy provides a bridge between microscopic dynamics and macroscopic irreversibility; included is a construction of quantum macroscopic states and a result on the equivalence of ensembles.

Contents

[*] On leave from Georgian National Astrophysical Observatory, Kazbegi ave. 2a, 0160 Tbilisi, Georgia.

R. Kotecký (ed.), *Methods of Contemporary Mathematical Statistical Physics*, Lecture Notes in Mathematics 1970, DOI: 10.1007/978-3-540-92796-9_6,

Part I. Fluctuations in Stochastic Lattice Gases

1 Introduction

One good way to learn about possible constructions of nonequilibrium statistical mechanics proceeds via the study of simple model systems. Traditionally, the so called stochastic lattice gases are playing there a prime role. A very early example is the Ehrenfest model. Not only was it important as one of the many urn models illustrating strategies and results from probability theory and from statistics, but it remains useful for learning about relaxation to equilibrium and about detailed balance, see e.g. [29] where Mark Kac does not seem to hesitate in calling the Ehrenfest model one of the most important models in all of physics. We will encounter that Ehrenfest model in the next section.

Over the last decades, many other lattice gas models have been invented. Often they obey simple updating rules and they are rather easy to visualize and to simulate (at least today). Yet, their behavior is rich, including sometimes clear examples of emergent behavior. The latter refers to the organization of robust structures or patterns, of critical behavior, and of phase transitions, which result from some collective or cooperative behavior between the many interacting components. They have appeared in interdisciplinary contexts, varying from models of traffic, to models of turbulence or to models for the spreading of infections, in computer science, in economy etc. On the more mathematics side, we have here an interesting ground for exploring and extending the theory of spatially extended Markov processes. The role of the spatial architecture of the processes has recently been more in the center of attention, e.g. in discussions of processes on random graphs, small worlds etc. In recent versions, the architecture (or graph) also undergoes a dynamics, in interaction with the particles. Clearly these lattice gas models have proven their use already. A few books with mathematical and statistical mechanical introductions to the theory of lattice gases include [30, 36, 56].

Many good references link stochastic lattice gases with fundamental problems in physics, be it in the context of turbulence or in the

derivation of hydrodynamic equations. In the present notes, we bring together a number of results in the construction of nonequilibrium statistical mechanics, as they appear for some simple stochastic lattice gas. The emphasis will be mostly on formal relations, from which both mathematical and physics treatments can find inspirations. We hope that this provides some step in a wider understanding of nonequilibrium issues in other and more realistic models. Indeed, one should remain aware that stochastic lattice gases are often only effective tools. They are Markovian from the start and their transition rates depend on some *ad hoc* choices. They are stochastic and there is no specification of a larger environment.

After a short reminder of aspects of a Markov dynamics for one particle, we introduce the main models in Section 3. That is continued in Section 4 where the steady state is further specified. In all that we work with a finite one-dimensional system on which there is a particle hopping in the bulk of the system and a birth and death process at the two boundary sites. We consider a time-dependent version of the dynamics in Section 5. The main tool is provided by a Lagrangian set-up in which a Girsanov formula specifies the action (Section 6). Section 7 gives the main fluctuation relations in the form of a Jarzynski identity (relating the irreversible work with the change in free energy) and of a fluctuation symmetry in the particle current (a so called steady state fluctuation theorem). While the results are mainly well-known we are not aware of a similar *unifying* presentation in the literature. There remain however very many nonequilibrium issues which are not treated in these notes. Some remarks are devoted to them in the final section, where we also describe the more general programme of dynamical fluctuations theory.

2 One Walker

Consider the set $K = \{0, \ldots, N\}$ and the discrete time Markov chain with transition probabilities $p(x, x') = x/N$ if $x' = x - 1$ and $p(x, x') = 1 - x/N$ if $x' = x + 1$ for $x, x' \in K$. One interpretation is to think of the state $x \in K$ as the number of particles in one of two vessels. The total number of particles over the two vessels is fixed equal to N. At each discrete time moment, one of the N particles is randomly selected and moved to the other vessel from where it was. That model is known as the Ehrenfest model (or dog-flea model).

As a mathematical object we have here an irreducible (but not aperiodic!) Markov chain (x_n) that satisfies the condition of detailed balance with respect to the stationary measure

$$\rho(x) = 2^{-N} \frac{N!}{x!(N-x)!}, \qquad x \in K$$

That condition of detailed balance

$$p(x,x')\ \rho(x) = p(x',x)\ \rho(x')$$

expresses the time-reversibility of the stationary process. Indeed, let $\boldsymbol{P}_\rho$ denote the stationary process on $K^{\mathbb{Z}}$ and define $y_n = x_{-n}$. The process (y_n) is Markovian with stationary law ρ. Its law is denoted by $\boldsymbol{P}_\rho\Theta$ where Θ stands for time-reversal. We show that $\boldsymbol{P}_\rho\Theta = \boldsymbol{P}_\rho$ as a consequence of the condition of detailed balance. The basic observation is that the transition probability for the process (y_n) is via Bayes' formula

$$q(y,y') = \text{Prob}[y_{n+1} = y' \,|\, y_n = y] = p(y',y)\,\frac{\rho(y')}{\rho(y)} = p(y,y')$$

Therefore, the condition of detailed balance is equivalent with the time-reversibility.

There is an easy way to generalize the above set-up. Let us first make the step to continuous time. We are now speaking about rates $c(x,y) \geq 0$ (or, transition probabilities per unit time) for the transition $x \to y$. If we assume that

$$c(x,y) = a(x,y)\, e^{[V(x)-V(y)]/2}$$

where $a(x,y) = a(y,x)$ is symmetric, then still

$$\frac{c(x,y)}{c(y,x)} = e^{V(x)-V(y)} \tag{2.1}$$

and $\rho(x) \propto \exp[-V(x)]$ is a reversible measure. It is remarkable that there remains a freedom in the choice of the symmetric function $a(x,y)$.

A new interpretation arises when thinking of the set $\{0,1,\ldots,N\}$ as the sites of a lattice interval, with the usual nearest neighbor connections. The rates $c(x,y)$ could be taken non-zero only if $y = x \pm 1$ in which case we have a nearest neighbor walk. The condition of detailed balance (2.1) assures that the walker will not drift; there is a potential landscape $V(x)$, $x = 0,\ldots,N$ which could be periodically repeated to cover all of $\mathbb{Z}$ if wished.

There are ways to break detailed balance. One could for example insert a non-zero transition rate for moving between the states (sites) $0 \leftrightarrow N$, and then take $c(x,x+1) = p$, $c(x,x-1) = q \neq p$, $N+1 \equiv 0$. In that case, say with $p > q$ there is a drift that the particle moves more

$x \to x+1$ than $x \to x-1$; there appears a net current. More generally, we can think of parameterizing the rates via

$$c(x,y) = a(x,y)\, e^{[V(x)-V(y)]/2}\, e^{s(x,y)/2}$$

where $s(x,y) = -s(y,x)$ would be antisymmetric. Of course, that does not yield a unique decomposition as we can e.g. put the difference $V(x)-V(y)$ inside $s(x,y)$. On the other hand, the $s(x,y)$ is of the form $s(x,y) = U(x) - U(y)$ if and only if it satisfies a gradient condition, that $s(x,y) + s(y,z) + s(z,x) = 0$ for all triples (x,y,z). The more important point however is that this term $s(x,y)$ has often a relevant physical interpretation. In what follows we will see it related to the entropy production. The entropy production is a physical notion that has arisen within irreversible thermodynamics, see e.g [21]. It goes well with considerations close to equilibrium.

The influence of the time-symmetric factor $a(x,y)$ is less understood.

The following sections will study some of the aspects above for multi-particle models. We now have a (possibly variable) number of particles and they move on the lattice following certain hopping rules. The effect of having many particles can result in (simpler) hydrodynamic behavior for macroscopic variables such as the density profile, but we will concentrate on the fluctuations instead.

3 Stochastic Lattice Gases

We start with a description of what is typically involved in stochastic lattice gases. We do not give the most general definitions but we specify to one special class.

3.1 States

By a lattice gas we understand a collection of particles whose positions are confined to the sites of a lattice. In some models the particles still have a momentum, most often with a finite number of possible values. Or, the particles can have extra decorations such as color or spin. Here we do not consider that.[1] The system thus consists of identical particles that can jump from site to site on the given architecture. The states of the system are assignments to each site of the number of particles.

[1] An important ingredient of the Hamiltonian (or symplectic) structure is thus lost. In particular the kinematical time-reversal that would normally change the sign of the velocities is absent.

To be specific we consider the finite linear chain $\Lambda_N = \{-N, -N+1, \ldots, 0, 1, \ldots, N-1, N\}$. The endpoints $i = \pm N$ in Λ_N will play a special role in what follows; we call them the boundary of the system while the other sites are in the bulk. Two sites i, j are nearest neighbors when $j = i \pm 1$.

We allow at most one particle per site i. We say that site i can be vacant or occupied. The state space (or the configuration space) is the finite set $K = \{0,1\}^{\Lambda_N}$. Elements of K are denoted by $\eta, \eta', \xi, \ldots$ and we write $\eta(i) \in \{0,1\}$ for the occupation at site $i \in \Lambda_N$.

3.2 Energy, Entropy, and Particle Number

One imagines a function H typically referred to as the Hamiltonian of the system,[2] that measures the energy of the state η. There is a great freedom of choice and all depends on the context or on the specific purpose. It does not hurt however to suppose something specific, say an energy function consisting of two terms:

$$H(\eta) = -B \sum_{i=-N}^{N} \eta(i) - \kappa \sum_{i=-N}^{N-1} \eta(i)\,\eta(i+1), \tag{3.1}$$

where B and κ are some constants. The first term contributes an energy $-B$ per particle being present in the system and the second term takes into account some form of nearest neighbor interaction related to the relative concentration of particles on neighboring sites.

Speaking of energy reminds us of its conservation law. We can indeed imagine that our system is in thermal contact with a very large heat bath at inverse temperature β (Boltzmann's constant is set equal to one), and for which all relevant changes are determined by the transitions in the system. In particular, every change $H(\eta') - H(\eta)$ in energy of the system is accompanied with the opposite change $\Delta E(\eta, \eta') = -(H(\eta') - H(\eta))$ of energy in the bath.[3] Imagining that

[2] It is of course not a Hamiltonian in the strict sense of analytical mechanics. The word Hamiltonian refers here more to the quantum world where one considers for example the hopping of electrons in a crystal structure. A mathematically precise correspondence, also for the dynamical properties, can often be achieved via the so called weak coupling limit or within the framework of Fermi's Golden Rule.

[3] In a more microscopic set-up, including a description of the degrees of freedom in the heat bath, one would need to specify a more exact decomposition of the total energy into the system part and the part that belongs to the reservoir. There would also be interaction terms, the coupling, that contain both system and reservoir variables. Obviously, some convention is then needed of what is system and what is reservoir variable.

the energy change of the reservoir is thermodynamically reversible, we associate to it a change of entropy (or an *entropy production*) in the reservoir equal to

$$\Delta S_{\rm res} = \beta \Delta E(\eta, \eta') = -\beta(H(\eta') - H(\eta))$$

In other words, every change $\eta \to \eta'$ in the system's configuration entails an entropy flux, that is β times the heat dissipated in the thermal reservoir.[4] For equilibrium purposes with just one heat bath, the relevant thermodynamic potential is the Helmholtz free energy. Its statistical mechanical version is

$$F = -\frac{1}{\beta} \log Z\,, \qquad Z = \sum_{\eta \in K} e^{-\beta H(\eta)}$$

Observe that if we change some parameter in H, e.g. the coupling coefficient κ in (3.1) (for fixed temperature), then the change in free energy $F = F(\kappa)$ equals the expected change in energy:

$$\frac{{\rm d}F}{{\rm d}\kappa} = \Big\langle \frac{{\rm d}H}{{\rm d}\kappa} \Big\rangle, \qquad H = H_\kappa$$

where

$$\langle g \rangle = \frac{1}{Z} \sum_{\eta \in K} g(\eta)\, e^{-\beta H(\eta)}$$

is the thermal expectation.

Another important observable is the particle number. We write

$$\mathcal{N}_{[j,k]}(\eta) = \sum_{i=j}^{k} \eta(i)$$

for the total number of particles in the lattice interval $[j,k] \cap \Lambda_N$, $-N \le j \le k \le N$. By construction, here we have that the particle numbers $\mathcal{N}_{[j,k]} \le 2N+1$ are *a priori* uniformly bounded. The total number of particles is denoted by $\mathcal{N} = \mathcal{N}_{[-N,N]}$. Making the correspondence with a gas again makes us think of a conservation law, now of the total number of particles. In what follows, we imagine that the system is also in contact with a particle reservoir at its boundary. Through the

[4] On a scale where one supposes that the η give the full microscopic description of the system, there is no associated change of entropy in the system. The total change of entropy (also called, the entropy production) is then also equal to $-\beta(H(\eta') - H(\eta))$. Most of the time however, there is a further *lower* level of description of the system variables with an associated degeneracy.

endpoints $i = \pm N$ particles can enter or leave the system. We can also speak of a birth or a death of a particle at these sites. In Section 4.2 we will introduce the particle currents. As particles can carry energy (see e.g. the first term in (3.1)), the flow of particles in and out of the system can also contribute to the change of energy in the reservoir, and hence to changes in entropy.

The equilibrium ensemble that allows both the exchange of energy and of particles is the grand-canonical one. It gives probabilities

$$\mathbb{P}^{\beta,a}[\eta] = \frac{1}{\mathcal{Z}}\, e^{a \sum \eta(i)}\, e^{-\beta H(\eta)} \tag{3.2}$$

where $\mathcal{Z} = \mathcal{Z}(a, \beta, N)$ is a normalization factor. The constant a is called the chemical potential (which is the standard term up to a factor β^{-1}) and in equilibrium it refers to and it is determined by the concentration of particles in the (imagined very large) environment.

3.3 Dynamics

The dynamics is given by a continuous time Markov process on K. We distinguish two modes of updating:

- A particle can jump (or hop) to nearest neighbor sites. That is a diffusion mechanism. We will not add external fields to the dynamics not to impose a bulk drift or bias;
- Particles can leave or enter the system at the boundary. That is a reaction mechanism. The system will be boundary driven.

We introduce some further notation to formalize the dynamics. As we only consider symmetric hopping, it is useful to introduce the transformation

$$\eta^{i,j}(k) = \begin{cases} \eta(k) & \text{if } k \neq i,\, k \neq j; \\ \eta(i) & \text{if } k = j; \\ \eta(j) & \text{if } k = i \end{cases} \tag{3.3}$$

That defines the configuration obtained from η after switching the occupation of the sites i, j. We allow only the hopping of particles to neighboring sites $j = i \pm 1$. The rate of the transition due to that diffusion mechanism is taken as

$$C(i, j, \eta) = \exp\Big[-\frac{\beta}{2}(H(\eta^{i,j}) - H(\eta))\Big], \quad |i - j| = 1 \tag{3.4}$$

where $H(\eta^{i,j})$ is the energy function after the transition and $H(\eta)$ is that corresponding to the initial configuration.

Analogously, we define the rate of birth and death $\eta \to \eta^i$ of the particles as

$$C(i,\eta) = e^{-a_i\eta(i)} \exp\Big[-\frac{\beta}{2}(H(\eta^i) - H(\eta))\Big] \tag{3.5}$$

where $H(\eta^i)$ is the energy function after the transition to η^i, the new configuration after the birth or the death of a particle at site i:

$$\eta^i(k) = \begin{cases} 1-\eta(k) & \text{if } k = i \\ \eta(k) & \text{if } k \neq i \end{cases} \tag{3.6}$$

To be definite we take births and deaths only at the boundary sites $i = -N, N$.

The physical interpretation of the dynamics is quite simple. Think of a one-dimensional channel in which particles diffuse while they enter or leave the system at its boundary. A biophysical realization seems to be found in the physics of ion channels connecting the inside and the outside of a living cell. The channel is a sort of opening or gate in the cell's membrane through which charged particles can move from higher to lower concentration, or following the gradient in electric potential etc. Here the relevant parameters are the values $a_{\pm N}$ which in fact represent the (different) chemical potentials of the two reservoirs at the outer edges.

With these definitions the Master equation governing the temporal behavior of probability measures on K is given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}_t(\eta) = &\sum_{i=1}^{N-1} [C(i,i+1,\eta^{i,i+1})\mathbb{P}_t(\eta^{i,i+1}) - C(i,i+1,\eta)\mathbb{P}_t(\eta)] \\ &+ C(-N,\eta^{-N})\mathbb{P}_t(\eta^{-N}) - C(-N,\eta)\mathbb{P}_t(\eta) \\ &+ C(N,\eta^{N})\mathbb{P}_t(\eta^{N}) - C(N,\eta)\mathbb{P}_t(\eta) \end{aligned} \tag{3.7}$$

That equation shows how the probability to find a given configuration in the system evolves in time. Alternatively, the generator L is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle f(\eta_t)\rangle = \langle Lf(\eta_t)\rangle \tag{3.8}$$

for functions ("observables") f on K, and where $\langle\cdot\rangle$ takes the expectation over the Markov process, including some (as yet unspecified) initial distribution. Explicitly,

$$Lf(\eta) = \sum_{i=1}^{N} C(i,i+1,\eta)[f(\eta^{i,i+1}) - f(\eta)] + C(-N,\eta)[f(\eta^{-N}) - f(\eta)] + C(N,\eta)[f(\eta^{N}) - f(\eta)] \quad (3.9)$$

Let us make a simple exercise by plugging in $f(\eta) = \eta(i)$ for some fixed i and by taking $\beta = 0$ in (3.4)–(3.5). The corresponding evolution equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \eta_t(i)\rangle = \langle \eta_t(i-1) + \eta_t(i+1) - 2\eta_t(i)\rangle$$

when $i \neq -N, N$, while for $i = \pm N$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \eta_t(i)\rangle = \langle \eta_t(i \mp 1) - \eta_t(i) + e^{-a_i\eta_t(i)}[1 - 2\eta_t(i)]\rangle$$

Apparently, these equations are closed in the density variables $\langle \eta_t(i)\rangle$, $i \in \Lambda_N$. In particular, putting their left-hand sides equal to zero, we get the stationary value $\langle \eta(i)\rangle = Ci + D$ for some constants C and D that depend on N and on the values $a_{\pm N}$. One checks that $a_{-N} = a_N = a$ if and only if $C = 0$, $D = 1/(1 + \exp(-a))$. When $C \neq 0$, then there is a linear density profile with slope $\sim 1/N$. Obviously, when repeating that calculation for $\beta \neq 0$, we run into a difficulty: the equation for the $\langle \eta_t(i)\rangle$ is no longer closed but there is a coupling with higher order correlation functions such as $\langle \eta_t(i)\,\eta_t(i+1)\rangle$. That feature is very generally true and it implies that we cannot simply solve the equations.[5] The stationary distribution is in general only implicitly known, as solution of the (time-independent) Master equation (3.7) with the left-hand side set zero.

3.4 Path-space Measure

One has to remember that a Markov process is a special probability distribution on paths. In our case, we have piecewise-constant paths. A path ω over the time-interval $[0,\tau]$ starts from an initial configuration η_0 after which it changes into $\eta_{t_1}, \eta_{t_2}, \ldots$ at random times $t_1, t_2, \ldots$ To be more precise we must add what is the configuration at the jump times as well. That is just a convention, and we take it

[5] The problem appears in all nontrivial dynamics for many particle systems. The resulting hierarchy of equations is sometimes referred to as the BBGKY-hierarchy, referring in particular to the hierarchy of equations that appear in kinetic gas theory for the various particle distribution functions. The study of possible ways of closing the hierarchy is a major concern in nonequilibrium physics.

that $\eta_{t_{k-1}} = \eta_{t_k-}$ and $\eta_{t_k} = \eta_{t_k+}$, or, the step-function is continuous from the right. An important transformation on path-space concerns the so called time-reversal Θ in which $(\Theta\omega)_t = \omega_{\tau-t}$, up to irrelevant modifications at the jump times making $\Theta\omega$ again right-continuous.

The random times are called the jump times of the process. The Markov process assigns a probability law to these times and to the corresponding transitions. There are two ingredients: the waiting time and the transition step. The waiting time is exponentially distributed with a weight $\lambda(\eta)$ that depends on the present configuration η. That waiting time is directly (and inversely) related to the escape rate

$$\lambda(\eta) = \sum_{\eta'} W(\eta \to \eta')$$

We will use the notation $W(\eta \to \eta')$ when indicating one allowed but general transition rate. The second ingredient sits in the transition rates as we have them introduced before. When the waiting time is over, a new configuration is chosen so that for time $t \downarrow 0$,

$$\text{Prob}[\eta_t = \eta' \,|\, \eta_0 = \eta] = (1 - \lambda(\eta)\, t)\, \delta_{\eta,\eta'} + W(\eta \to \eta')\, t + o(t)$$

A more explicit realization of that path-space measure goes via Girsanov's formula, see Section 6.

4 Steady State

4.1 Detailed Balance

One observes from the definition (3.4) that:

$$\frac{C(i,j,\eta)}{C(i,j,\eta^{i,j})} = \frac{\exp\left[-\beta H\left(\eta^{i,j}\right)\right]}{\exp\left[-\beta H\left(\eta\right)\right]} = \frac{\mathbb{P}^{\beta,a}[\eta^{i,j}]}{\mathbb{P}^{\beta,a}[\eta]} \tag{4.1}$$

where we have inserted the ratio of probabilities according to (3.2). That is verified for all values a. Furthermore, with the definition (3.5) we have

$$\frac{C(i,\eta)}{C(i,\eta^i)} = \frac{\exp[-a_i\eta(i)]}{\exp[-a_i(1-\eta(i))]} \frac{\exp\left[-\beta H\left(\eta^i\right)\right]}{\exp\left[-\beta H\left(\eta\right)\right]}, \quad i = \pm N \tag{4.2}$$

Comparing with formula (3.2), still for $i = \pm N$ and for $a_{-N} = a_N = a$,

$$\frac{C(i,\eta)}{C(i,\eta^i)} = \frac{\mathbb{P}^{\beta,a}[\eta^i]}{\mathbb{P}^{\beta,a}[\eta]} \tag{4.3}$$

Summarizing, when the particle reservoirs left and right have equal concentration, then the system dynamics satisfies the condition of detailed balance

$$\frac{W(\eta \to \eta')}{W(\eta' \to \eta)} = \frac{\mathbb{P}^{\beta,a}[\eta']}{\mathbb{P}^{\beta,a}[\eta]} \tag{4.4}$$

for all allowed transitions $\eta \to \eta'$ and corresponding transition rates $W(\eta \to \eta')$. Under that same condition $a_{-N} = a_N = a$ we thus have that (3.2) is a reversible stationary measure. The corresponding process is the steady state for equilibrium conditions.

Observe that if we consider unequal rates at the boundaries $a_1 \neq a_N$ then we could still try

$$\mathbb{P}^{\beta,a_1,a_N}(\eta) = \frac{1}{Z} \exp[-\beta H(\eta)] \exp[a_1\eta(1) + a_N\eta(N)] \tag{4.5}$$

as a candidate stationary distribution. In that case the analogue of (4.3) is still verified. Yet, the condition (4.1) now fails.

4.2 Nonequilibrium Model

Now comes the question what happens when $a_1 \neq a_N$. Let us first consider the left boundary of the system, for which we can write, see (4.2),

$$\frac{C(-N,\eta)}{C(-N,\eta^{-N})} = e^{-\beta[H(\eta^{-N})-H(\eta)]-a_{-N}J_\ell(\eta,\eta^{-N})} \tag{4.6}$$

where $J_\ell(\eta,\eta^{-N}) = 1$ when the particle leaves the system via the site $-N$, i.e., $\eta(-N) = 1$, and $J_\ell(\eta,\eta^{-N}) = -1$ when a new particle enters, i.e., $\eta(-N) = 0$. That is an antisymmetric current of particles, taking positive when the particles leave the system. In the same way we define the current $J_r(\eta,\eta') = 1$ when $\eta(N) = 1$, $\eta' = \eta^N$ and $J_\ell(\eta,\eta') = -1$ when $\eta(N) = 0$, $\eta' = \eta^N$. The currents are zero otherwise.

Taking all transitions together, we have

$$\frac{W(\eta \to \eta')}{W(\eta' \to \eta))} = e^{-\beta[H(\eta')-H(\eta)]-a_{-N}J_\ell(\eta,\eta')-a_N J_r(\eta,\eta')} \tag{4.7}$$

One recognizes the change of entropy in the environment:

$$S(\eta,\eta') = \beta\Delta E(\eta,\eta') - \mu_\ell \Delta\mathcal{N}_\ell(\eta,\eta') - \mu_r \Delta\mathcal{N}_r(\eta,\eta')$$

where $\mu_\ell = a_{-N}$ respectively $\mu_r = a_N$ are the chemical potentials (up to some factor β that we have ignored) of the particle reservoirs left and right, and $J_\ell = \Delta\mathcal{N}_\ell$, $J_r = \Delta\mathcal{N}_r$ are the changes in particle number in the *left*, respectively *right* particle reservoir. The form (4.7) or

$$\frac{W(\eta\to\eta')}{W(\eta'\to\eta)} = e^{S(\eta,\eta')}$$

is known as that of local detailed balance.[6]

The currents J_ℓ and J_r appear in the conservation law for the particle number. The sum of these currents equals the number of particles that leave the system,

$$J_\ell(\eta,\eta') + J_r(\eta,\eta') = \mathcal{N}(\eta) - \mathcal{N}(\eta') \tag{4.8}$$

or

$$a_{-N}J_\ell(\eta,\eta') + a_N J_r(\eta,\eta') = (a_{-N}-a_N)J_\ell + a_N(\mathcal{N}(\eta)-\mathcal{N}(\eta')) \tag{4.9}$$

From now on, we write $a_N = a$, $a_{-N} = a+\delta$ so that

$$\frac{W(\eta\to\eta')}{W(\eta'\to\eta)} = \frac{\mathbb{P}^{\beta,a}[\eta']}{\mathbb{P}^{\beta,a}[\eta]}\, e^{-\delta J_\ell(\eta,\eta')}$$

The parameter δ thus measures some distance to the equilibrium situation, and enables the tentative terminology of *close* versus *far* from equilibrium.

As above we define the bulk currents $J_i(\eta,\eta')$ to be $+1$ if in the transition $\eta\to\eta'$ a particle moves over the bond $i\to i+1$, and equal to -1 if a particle moves $i\leftarrow i+1$. More generally,[7] we consider a path $\omega = (\eta_t)_{t=0}^{\tau}$ and currents $J_i(\omega)$, $i=-N,\ldots,N-1$, defined by

$$J_i(\omega) = J_i(\eta_0,\eta_{t_1}) + J_i(\eta_{t_1},\eta_{t_2}) + \ldots + J_i(\eta_{t_{n-1}},\eta_\tau)$$

[6] Remark that in (4.7) a possible time-symmetric prefactor to the rates (3.5) or (3.4) will never appear; there is only the part that is antisymmetric under $\eta\leftrightarrow\eta'$. The fact that the entropy production appears as the source term of the breaking of time-reversal symmetry, is no accident but it is related to more general considerations that here are simply applied in order to obtain a reasonable physical interpretation of our effective dynamics, see e.g. [38, 40].

[7] In fact and throughout we call current what is more like a time-integrated current, or a change of particle number.

In particular, $J_r = J_N$ and for $i < k$,

$$J_i(\omega) - J_k(\omega) = \mathcal{N}_{[i+1,k]}(\eta_\tau) - \mathcal{N}_{[i+1,k]}(\eta_0)$$
$$J_\ell(\omega) + J_{-N}(\omega) = \eta_0(-N) - \eta_\tau(-N) \tag{4.10}$$

Observe that the currents J_i are extensive in the time τ.

All of that is related to the process, be it transient or be it steady. Except for the following section however, we will be mostly interested in the steady state regime. It is easy to verify that we have here a unique stationary distribution ρ. It satisfies the time-independent Master equation (3.7) (zero left-hand side). Corresponding to ρ there is then a stationary process with distribution $\boldsymbol{P}_\rho$. If we look at expectations in the stationary process we write $\langle\,\cdot\,\rangle_\rho$. From the conservation laws (4.8) and (4.10) we have

$$\langle J_\ell\rangle_\rho = -\langle J_r\rangle_\rho = -\langle J_i\rangle_\rho, \qquad i = -N, \dots, N-1$$

There are alternative expressions for these expectations by using the dynamical equations (3.8). For example,

$$\frac{1}{\tau}\langle J_i\rangle_\rho = \langle C(i,i+1,\eta)(\eta(i) - \eta(i+1)\rangle_\rho = \langle 1 - \eta(N)[1 + e^{-a}]\rangle_\rho$$

5 Time-dependent Dynamics

5.1 Modifications with respect to Section 4

Nonequilibrium conditions can be obtained in a variety of ways. The above gives a set-up for boundary driven steady states. Another way of driving the system away from equilibrium is by applying an external *bulk* field. We consider here a modification which also frustrates the system (as it cannot simply relax to equilibrium). We remain with the same states but the updating becomes time-dependent. The idea is that the values of parameters in the Hamiltonian are changed while the dynamics enrolls.

We have a time-dependent Hamiltonian H_t so that the transition rates $W_t(\eta \to \eta')$ are also depending on the moment t of the jump $\eta \to \eta'$. For example, the rate for exchanging the occupation at sites i and $j = i \pm 1$ is

$$C_t(i,j,\eta) = \exp\Big[-\frac{\beta}{2}\big(H_t(\eta^{i,j}) - H_t(\eta)\big)\Big], \qquad |i-j| = 1 \tag{5.1}$$

(compare with (3.4)) depending on the time t.

There is no longer a very good sense in which we can speak about the stationary distribution. Still we can consider for each H_t, $t \in [0,\tau]$, the corresponding Gibbs distribution

$$\rho_t(\eta) = \frac{1}{\mathcal{Z}_t}\, e^{a\sum \eta(i)}\, e^{-\beta H_t(\eta)} \tag{5.2}$$

where $\mathcal{Z}_t = \mathcal{Z}_t(a,\beta,N)$ is now also time-dependent. There is an associated free energy

$$\mathcal{A}_t = -\frac{1}{\beta}\log \mathcal{Z}_t \tag{5.3}$$

In the time-dependent case, we will only work with the dynamics for which $\delta = 0$, $a_{-N} = a_N = a$ fixed, i.e., there is just one particle reservoir and one heat bath reservoir.[8]

5.2 Work and Heat

When there is an energy exchange between system and reservoir, there is heat. For a history $\omega = (\eta_t)_0^\tau$ where the jumps in the configuration happen at times $t_1, t_2, \ldots t_n$, the total heat $\mathcal{Q}$ transferred to the system is the sum of differences of energy:

$$\begin{aligned}\mathcal{Q} = H_{t_1}(\eta_{t_1}) - H_{t_1}(\eta_0) + H_{t_2}(\eta_{t_2}) - H_{t_2}(\eta_{t_1}) + \ldots \\ + H_{t_n}(\eta_\tau) - H_{t_n}(\eta_{t_{n-1}})\end{aligned} \tag{5.4}$$

On the other hand, the work $\mathcal{W}$ performed upon the system is a sum of changes of the Hamiltonian at fixed configurations:

$$\begin{aligned}\mathcal{W} = H_\tau(\eta_\tau) - H_{t_n}(\eta_\tau) + H_{t_n}(\eta_{t_{n-1}}) - H_{t_{n-1}}(\eta_{t_{n-1}}) + \ldots \\ + H_{t_1}(\eta_0) - H_0(\eta_0)\end{aligned} \tag{5.5}$$

Therefore, as an expression of the first law of thermodynamics,

$$\mathcal{Q} + \mathcal{W} = H_\tau(\eta_\tau) - H_0(\eta_0) \tag{5.6}$$

is the total change of system energy between the initial and the final configurations η_0 and η_τ in the path ω.

[8] We can of course make a dynamics such that at every moment the distribution $\mathbb{P}_t = \rho_t$ exactly coincides with (5.2). We could e.g. take a transition rate $\bar{W}_t(\eta \to \eta') = \rho_t(\eta')$. One can think of it as admitting an infinitely fast relaxation of the equilibrium process. Alternatively, one can think of an ultra-slow time-dependence in H_t so that, before any change, the system has already relaxed to the equilibrium distribution corresponding to the instantaneous value.

6 Lagrangian Set-up: Girsanov Formula

As we have seen in the course of our computation around and below (3.9), the evolution equations give a hierarchy of equations for the various correlation functions of the stationary distribution. Solving them is like diagonalizing a large matrix[9] and it is not even clear whether it would always permit us to extract the most relevant information. A more global characterization of the stationary distribution is perhaps obtained by going to a space-time picture. On that level the process is space-time local and explicit. The variables are the histories or trajectories of the system. It brings us in line with a Lagrangian approach to statistical mechanics, [38], as pioneered by Onsager and Machlup.

Given two Markov processes on the same space K, we can consider two path-space measures $\boldsymbol{P}$ and $\bar{\boldsymbol{P}}$ with corresponding escape rates λ and $\bar{\lambda}$, and transition rates W and $\bar{W}$. We consider all paths on the interval $[0, \tau]$ and we assume that for all η,

$$\{\eta', W(\eta \to \eta') \neq 0\} = \{\eta', \bar{W}(\eta \to \eta') \neq 0\}$$

We can then look for the density of $\boldsymbol{P}$ with respect to $\bar{\boldsymbol{P}}$. That density is a Radon-Nikodym derivative and can be written down quite explicitly in the so called Girsanov formula:

$$\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}\bar{\boldsymbol{P}}}(\omega) = \exp\Big[-\int_0^\tau \big(\lambda(\eta_t) - \bar{\lambda}(\eta_t)\big)\,\mathrm{d}t + \sum_{t\leq\tau} \log\frac{W(\eta_{t-} \to \eta_t)}{\bar{W}(\eta_{t-} \to \eta_t)}\Big] \quad (6.1)$$

when restricted to events that are measurable from the trajectory in $[0, \tau]$. The last sum in the exponential is over the jump times, as they appear in the path ω. We have assumed here that the two processes $\boldsymbol{P}$ and $\bar{\boldsymbol{P}}$ start from the same configuration. If they have different initial distributions ν and $\bar{\nu}$, then a prefactor $\nu(\eta)/\bar{\nu}(\eta)$, $\omega_0 = \eta$, must be added to the right-hand side of the Girsanov formula (6.1). The formula remains intact when the process is not time-homogeneous. One then adds the correct time-dependence to the escape and to the transition rates.[10]

[9] In fact, the generator of a Markov dynamics need not even be diagonalizable. That is a possible consequence of the breaking of detailed balance.

[10] The formula is also not really restricted to Markov processes, or to finite state spaces. The more useful way of considering that formula is as a generalization of the Boltzmann-Gibbs formula (3.2), where the essential input is that one can make sense of what is written in the exponential as a sum of quasi-local terms. Here we can speak about the action as a sum of a local Lagrangian. Indeed, if we write out the rates $W(\eta \to \eta')$ of our stochastic lattice gas, and we take $\bar{\boldsymbol{P}}$ say

Our application of the Girsanov formula will concern time-reversal. If we have a distribution $\boldsymbol{P}$ on paths, then its time-reversal $\boldsymbol{P}\Theta$ is obtained via

$$\frac{\mathrm{d}\boldsymbol{P}\Theta}{\mathrm{d}\boldsymbol{P}^0}(\omega) = \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}\boldsymbol{P}^0}(\Theta\omega)$$

for an arbitrary process $\boldsymbol{P}^0$ which is reversible. The dependence on initial configurations is again ignored, but it is essential in the consideration of the time-reversal invariant process $\boldsymbol{P}^0$.

7 Fluctuation Relations for the Entropy Production

7.1 Jarzynski Equality

Recall the set-up for the time-dependent dynamics in Section 5. We take the case where $a_1 = a_N = a$. The Jarzynksi identity is a relation between the work $\mathcal{W}$ of (5.5) and the change in free energy (5.3). In the context of stochastic lattice gases, we get it as

$$\mathbb{E}^a_{\rho_0}[e^{-\beta\mathcal{W}}] = e^{-\beta\Delta\mathcal{A}}, \tag{7.1}$$

The left-hand side is the expectation in the time-dependent dynamics with fixed chemical potential (left and right) equal to a and with Hamiltonian H_t at inverse temperature β, started from equilibrium ρ_0 at time $t = 0$. The right-hand side contains the difference

$$\Delta\mathcal{A} = \mathcal{A}_\tau - \mathcal{A}_0, \tag{7.2}$$

of free energies[11] (5.3).

---

corresponding to $\beta = 0$, $a_{-N} = a_N = 0$ we obtain there a sum over space-time of local interaction terms. Without trying to formalize the idea, see however [39], one can thus consider the stationary distribution to be the projection (or restriction) of that space-time path-space measure to an equal time layer, see e.g. [34]. There is no *a priori* reason why that projected measure should inherit a spatial locality, see e.g. [44].

[11] To make sure, there is no assumption that at any future time $t > 0$ (including at time τ) the distribution should be the ρ_t of (5.2). We are starting from equilibrium at time zero, but then the system is most likely away from instantaneous equilibrium with respect to the Hamiltonian H_t. Yet, the result of (7.1) is a statement about equilibrium free energies. We can measure these (and how they possibly depend on some parameter) via some nonequilibrium procedure.

Proof. To prove (7.1) we make a first application of the Girsanov formula. The two distributions correspond to our time-dependent process $\boldsymbol{P}^a_{\rho_0}$ on the one hand and to the time-reversed process $\bar{\boldsymbol{P}}^a_{\rho_\tau}\Theta$ on the other hand. By $\bar{\boldsymbol{P}}^a_{\rho_\tau}$ we mean the process started at time zero from the distribution ρ_τ and with time-reversed protocol, i.e., the rates are $\bar{W}_t = W_{\tau-t}$.

We have therefore for a fixed path ω with jump times $t_1, \ldots, t_n$ in the interval $[0, \tau]$, that

$$\frac{\mathrm{d}\boldsymbol{P}^a_{\rho_0}}{\mathrm{d}\bar{\boldsymbol{P}}^a_{\rho_\tau}\Theta}(\omega) = \frac{\rho_0(\omega_0)}{\rho_\tau(\omega_\tau)}\, \exp R(\omega) \tag{7.3}$$

with, from (6.1),

$$R = \frac{W_{t_1}(\eta_0 \to \eta_{t_1}) W_{t_2}(\eta_{t_1} \to \eta_{t_2}) \ldots W_{t_n}(\eta_{t_{n-1}} \to \eta_\tau)}{W_{t_n}(\eta_\tau \to \eta_{t_{n-1}}) \ldots W_{t_2}(\eta_{t_2} \to \eta_{t_1}) W_{t_1}(\eta_{t_1} \to \eta_0)} \tag{7.4}$$

By using the detailed balance relations

$$\frac{W_t(\eta \to \eta')}{W_t(\eta' \to \eta)} = \exp[-\beta(H_t(\eta') - H_t(\eta)) + a(\mathcal{N}(\eta') - \mathcal{N}(\eta))]$$

and combining that with the expression (5.4) for the heat, the ratio (7.4) reduces to

$$R(\omega) = e^{-\beta \mathcal{Q}}\, e^{a[\mathcal{N}(\eta_\tau) - \mathcal{N}(\eta_0)]}$$

Hence, looking back at (7.3) and substituting (5.2), (5.6) and (7.2), we have

$$\begin{aligned} \log \frac{\mathrm{d}\boldsymbol{P}^a_{\rho_0}}{\mathrm{d}\bar{\boldsymbol{P}}^a_{\rho_\tau}\Theta}(\omega) &= -\beta \mathcal{Q}(\omega) + \log \frac{\mathcal{Z}_\tau}{\mathcal{Z}_0} - \beta[H_0(\eta_0) - H_\tau(\eta_\tau)] \\ &= \beta\, [\mathcal{W}(\omega) - \Delta\mathcal{A}] \end{aligned} \tag{7.5}$$

The Jarzynski equality (7.1) is then an easy consequence of the normalization of path-space measures:

$$\int \mathrm{d}\boldsymbol{P}^a_{\rho_0}(\omega) \frac{\mathrm{d}\bar{\boldsymbol{P}}^a_{\rho_\tau}\Theta}{\mathrm{d}\boldsymbol{P}^a_{\rho_0}}(\omega) = 1 \tag{7.6}$$

For further background information on these relations between irreversible work and free energy differences, one can check e.g. [7, 25, 38].

7.2 The Direction of Particle Current

We come back to the time-homogeneous nonequilibrium process as we had it first in Section 4.2. Physically we expect that there will be a particle current flowing from higher to lower concentration. To be specific, let us assume that $\delta \geq 0$, $a_{-N} \geq a_N$, so that the physical picture suggests that the mean particle current $\langle J_i \rangle_\rho \geq 0$. The question is how to actually see that. Remember that we do not know a thing about the stationary distribution ρ in general. Nevertheless the direction of the particle current will easily follow within our set-up.

From the Girsanov formula (6.1) for $\boldsymbol{P}_\rho$ with respect to $\boldsymbol{P}_\rho\Theta$, both started in the stationary distribution ρ, we have

$$\frac{\mathrm{d}\boldsymbol{P}_\rho}{\mathrm{d}\boldsymbol{P}_\rho\Theta}(\omega) = \frac{\rho(\omega_0)}{\rho(\omega_\tau)} \exp\left[-\beta\left(H(\omega_\tau) - H(\omega_0)\right) + a\Delta\mathcal{N} - \delta J_\ell(\omega)\right], \tag{7.7}$$

Again from the normalization we have:

$$\int \mathrm{d}\boldsymbol{P}_\rho(\omega) \frac{\mathrm{d}\boldsymbol{P}_\rho\Theta}{\mathrm{d}\boldsymbol{P}_\rho}(\omega) = 1.$$

and hence, by concavity,

$$\int \mathrm{d}\boldsymbol{P}_\rho(\omega) \log \frac{\mathrm{d}\boldsymbol{P}_\rho\Theta}{\mathrm{d}\boldsymbol{P}_\rho}(\omega) \leq 0.$$

But, from (7.7) and by stationarity

$$0 \leq \int \mathrm{d}\boldsymbol{P}_\rho \log \frac{\mathrm{d}\boldsymbol{P}_\rho}{\mathrm{d}\boldsymbol{P}_\rho\Theta}(\omega) = -\delta\langle J_\ell \rangle_\rho = \delta\langle J_i \rangle_\rho \tag{7.8}$$

We conclude that

$$\delta\langle J_i \rangle_\rho \geq 0 \tag{7.9}$$

which shows that the average direction of the particle current depends only on the sign of δ. See [42] for a very similar analysis in the case of heat conduction. To get a strict inequality $\langle J_i \rangle_\rho > 0$ is also possible for $\delta > 0$; it suffices to see that there is a non-zero probability that the current J_i as a function of the path ω is not constant equal to zero even when $\omega_0 = \omega_\tau$.

7.3 Fluctuation Theorem

The previous results were all a direct consequence of the normalization condition applied to the Radon-Nikodym derivative (6.1) between two path-space measures. Here we go for a result that is somewhat more

detailed and concerns a symmetry in the fluctuations of the current. We follow the method of [38, 39, 43] but the present model can also be treated via [35].

We fix an $i = -N, \ldots, N-1$ and consider the current J_i as function of the path over the interval $[0, \tau]$. Define the generating function $q(\lambda)$, $\lambda \in \mathbb{R}$ by

$$q(\lambda) = \lim_{\tau \uparrow +\infty} \frac{1}{\tau} \log \langle e^{-\lambda J_i} \rangle_\rho \tag{7.10}$$

The limit exists by the Perron-Frobenius theorem, and is independent of $i = -N, \ldots, N-1$ because of (4.10). The fluctuation symmetry is that

$$q(\lambda) = q(\delta - \lambda) \tag{7.11}$$

Before providing a proof observe that $q(\lambda)$ is the Legendre transform of the rate functions of large deviations for J_i. The interested reader is referred to the literature on large deviations (and the Gärtner-Ellis theorem in particular) for more details, see e.g. [9]. The idea is that

$$\boldsymbol{P}_\rho[J_i \simeq \tau j] \simeq e^{-\tau I(j)}, \qquad \tau \uparrow +\infty \tag{7.12}$$

with

$$I(j) = \inf_\lambda (-\lambda j - q(\lambda)) \tag{7.13}$$

Substituting the identity (7.11), we get

$$I(j) = \inf_\lambda \big((-\delta + \lambda)(-j) - q(\delta - \lambda) \big) - \delta\, j = -\delta\, j + I(-j) \tag{7.14}$$

or

$$I(j) - I(-j) = -\delta\, j \tag{7.15}$$

That can be again translated to an identity for (7.12):

$$\frac{\boldsymbol{P}_\rho(J_i/\tau \simeq j)}{\boldsymbol{P}_\rho(J_i/\tau \simeq -j)} \simeq e^{\tau\, \delta\, j} \tag{7.16}$$

The interpretation of that expression is that there exists a relation between the probabilities of having a current $+j$ and $-j$; the probability of having a particle current in the direction opposite to the expected one ($\delta j > 0$) is exponentially small with $\tau \to +\infty$. The formula (7.16) has appeared before and many papers have been devoted to proving it in a variety of contexts. It first appeared in the context of smooth dynamical systems where it concerned the fluctuations of the phase space contraction, [17, 16, 54], and the result has become known as the (steady state Gallavotti-Cohen) fluctuation theorem. We now prove (7.11).

Proof. By definition and since $J_i(\Theta\omega) = -J_i(\omega)$,

$$\langle e^{-\lambda J_i}\rangle_\rho = \int \mathrm{d}\boldsymbol{P}_\rho(\Theta\omega)\, e^{\lambda J_i(\omega)}$$

Next we insert the Radon-Nikodym derivative

$$\int \mathrm{d}\boldsymbol{P}_\rho(\omega)\frac{\mathrm{d}\boldsymbol{P}_\rho\Theta}{\mathrm{d}\boldsymbol{P}_\rho}(\omega)\, e^{\lambda J_i(\omega)} = \int \mathrm{d}\boldsymbol{P}_\rho(\omega)\exp[\Delta - \delta J_i(\omega)]\, e^{\lambda J_i(\omega)} \tag{7.17}$$

where, via (6.1) or via (7.7),

$$\Delta = \log\frac{\rho(\omega_\tau)}{\rho(\omega_0)} + \beta\,(H(\omega_\tau) - H(\omega_0)) - a\Delta\mathcal{N} + \delta(J_\ell(\omega) + J_i(\omega)) \tag{7.18}$$

Notice that $|\Delta\mathcal{N}|=|\mathcal{N}(\omega_\tau)-\mathcal{N}(\omega_0)| \le 2N+1$. Further, $|\log\rho(\omega_\tau)/\rho(\omega_0)|$ is also bounded because at any rate, $\rho(\eta) \neq 0$, $\eta \in K$. Finally, there is the conservation law

$$J_\ell(\omega) + J_i(\omega) = -\mathcal{N}_{[-N,i]}(\omega_\tau) + \mathcal{N}_{[-N,i]}(\omega_0) \tag{7.19}$$

implying that the sum of these currents is also bounded, and $|H(\omega_\tau) - H(\omega_0)| \le 4(\kappa + B)N$, see (3.1). Hence we conclude that $|\Delta| \le$ const (with a constant that also depends on N but not on τ) which finishes the proof.

In the case $\beta = 0$ (the bulk dynamics is that of the so called simple symmetric exclusion process), more is known about the fluctuations of the current, see e.g. [13].

8 More Nonequilibrium Issues

Only a limited review has been given in the previous sections of recent work on nonequilibrium aspects of stochastic lattice gases. It basically involved dynamical fluctuations of time-antisymmetric quantities only, such as particle currents and entropy fluxes. We attempt to give some additional remarks.

8.1 Escape from Equilibrium

Imagine yourself enclosed in a well-isolated room. At time zero somebody opens doors and windows; how long would it take before you feel that? Probably you will become aware of the openings because of some air current, and its intensity will depend on the outdoor conditions.

In the present section we use our model to ask a similar question. Suppose a probability distribution μ on K which is nonzero, $\mu(\eta) > 0$, only when $\mathcal{N}(\eta) = m$ for some given number $0 < m < N$ of particles. For the rest we assume that it is thermally distributed:[12]

$$\mu(\eta) = \frac{1}{\mathcal{Z}_m} \exp[-\beta H(\eta)]\, \chi(\mathcal{N}(\eta) = m) \tag{8.1}$$

We consider the dynamics of Section 4.2 in the steady state $\boldsymbol{P}_\rho$ and we ask for the fraction of times that we see a fixed configuration η:

$$p_\tau(\eta) = \frac{1}{\tau} \int_0^\tau \chi(\eta_t = \eta)\, \mathrm{d}t$$

That time-average is a random variable and it depends on η. An important question in the theory of large deviations is to ask whether these fractions resemble a given probability measure; here we ask what is the function $I(\mu)$ so that

$$\boldsymbol{P}_\rho[p_\tau \simeq \mu] \simeq e^{-\tau\, I(\mu)}, \qquad \tau \uparrow +\infty \tag{8.2}$$

That question has been rigorously studied by Donsker and Varadhan, [9, 14], and we know an expression for $I(\mu)$:

$$I(\mu) = -\inf_{g>0} \Big\langle \frac{Lg}{g} \Big\rangle_\mu \tag{8.3}$$

The expectation is over the distribution μ and Lg is as before the generator of the process acting on a function g (over which we vary in (8.3)). It is known that in the case of a detailed balance process, the minimizer in (8.3) is $g^\star = \sqrt{\mu/\rho}$.

From (8.3) it is clear that we must find the function g on K that minimizes

$$\sum_{\eta\in K} \frac{\mu(\eta)}{g(\eta)} \Big\{ \sum_{i=1}^{N-1} C(i, i{+}1, \eta)\, g(\eta^{i,i+1}) + g(\eta^{-N})\, C(-N, \eta) + g(\eta^N)\, C(N, \eta) \Big\}$$

The sum is effectively over all η with $\mu(\eta) > 0$ or $\mathcal{N}(\eta) = m$. Since the configurations $\eta^{\pm N}$ have one particle more or less than η, we can put $g(\xi) = 0$ whenever $\mathcal{N}(\xi) \neq m$, and the minimization is over

$$\sum_{\eta\in K} \frac{\mu(\eta)}{g(\eta)} \sum_{i=1}^{N-1} C(i, i+1, \eta)\, g(\eta^{i,i+1})$$

[12] χ is the indicator function.

Taking again the complete Donsker-Varadhan functional we have

$$I(\mu) = \langle C(-N, \eta(-N)) + C(N, \eta(N)) \rangle_\mu - \inf_g \Big\langle \frac{\mathcal{L}g}{g} \Big\rangle_\mu \tag{8.4}$$

where the new generator $\mathcal{L}$ only considers particle exchanges generating a dynamics that satisfies the condition of detailed balance with respect to our μ; in other words, it is typical to 'see' the distribution μ for that pure hopping process. It follows that the second term in (8.4) is zero to conclude that

$$I(\mu) = \langle C(-N, \eta(-N)) + C(N, \eta(N)) \rangle_\mu$$

As we could expect the fluctuation functional measures the dynamical activity or intensity that lets the system escape from μ. There is in fact a more general relation between occupation statistics as we have it here and the time-symmetric notion of traffic or dynamical activity, as discussed in [45, 46]. Close to equilibrium that traffic becomes approximately equivalent with the entropy production rate. This observation lies behind the emergence of certain nonequilibrium variational principles, see [47].

The issue here is somewhat related to the problem of metastability, as found also in the contributions by Anton Bovier and by Frank den Hollander, see also [4]. At the same time, it is related to e.g. Section 5.7 in [9] (diffusion from a domain), and from a physics point of view it is related to Kramers' theory, [22].

8.2 Macroscopic Fluctuations

The previous calculation was to give a glimpse of the theory of dynamical fluctuations. In a more general context it discusses the statistical properties of empirical time-averaged quantities in both time-symmetric and time-antisymmetric sectors, including their correlations. There is however another scale of description on which similar questions become better manageable, that is the level of macroscopic fluctuations, static and dynamical. For our model, it refers to a hydrodynamical scaling in which one observes the evolution of the density profile. After a diffusive space-time rescaling one finds that the density profile $n_t(r)$ obeys a standard diffusion equation, [30],

$$\frac{\partial n_t(r)}{\partial t} = \frac{1}{2} \frac{\partial}{\partial r} D(n_t(r)) \frac{\partial}{\partial r} n_t(r)\,, \qquad r \in [-1, 1]$$

where $D(n_t(r))$ is the diffusion 'constant,' further constrained by imposing the boundary conditions $\rho(\pm 1) = 1/(1 + e^{a_\pm})$. In the simplest

case (corresponding to $\beta = 0$) the diffusion is truly constant and the stationary profile $n^\star$ is linear. One can however ask for fluctuations around that (typical) behavior. The hydrodynamic equation is the result of a law of large numbers and we can ask for the plausibility of a deviating density profile. We refer to [1, 13, 31, 41, 48] for further results and insights.

Part II. Macroscopic Irreversibility

9 Introduction

Up to now we have been discussing so called mesoscopic systems, or more precisely, classical mesoscopic systems modeled as stochastic processes. Time-reversal symmetry was broken by applying external conditions, frustrating the system in its return to equilibrium. However, the microscopic laws of nature are time-reversal invariant. One could then perhaps have expected to find that all resulting behavior is invariant under time-reversal, except perhaps for some microscopic interactions.[13] That is not what we see: systems return to equilibrium thereby showing the infamous arrow of time. The equations of macroscopic physics are not time-reversible (or not always). They have often been described and been used quite some time before their microscopic origin was clarified. In fact their (macroscopic) irreversibility once casted doubt on the kinetic and atomistic picture of matter and motion. One of the greatest successes in the pioneering days of statistical mechanics was then indeed the explanation of that manifest irreversibility.

That the emergent macroscopic laws are irreversible is not so difficult to understand at least qualitatively. One should realize that distinct macroscopic states can be very different in the number of microstates they consist of. It is the installation of an initial macrostate that breaks the invariance under time-reversal: unless forbidden by additional constraints, a less plausible initial state evolves to a more plausible macrostate and finally to the most plausible, called equilibrium, exactly because that is more plausible. The plausibility is measured in terms of the 'number' of microstates or, more precisely, by the Boltzmann or counting entropy which has a well defined thermodynamic limit (a precise meaning of that counting needs to be and will be specified). The generic increase of the entropy between initial and

[13] All would probably agree e.g. that the weak fundamental interaction is not at all responsible for macroscopic irreversibility, and most would probably agree that quantum mechanics is not either (while this is somewhat more tricky).

final macrostates (traditionally both in equilibrium) is known as the second law of thermodynamics and can be formulated in various ways. Still, even more is often true: considerations of entropy via counting the microstates consistent with a given macrostate, are *a priori* not restricted only to the initial and the final states and can be applied to each intermediate, 'nonequilibrium' state as well. Extended in that way, the Boltzmann entropy is often an increasing function of time, as was first demonstrated for the Boltzmann equation, the macroscopic evolution equation for rare gases, and rigorously proven in the so called Boltzmann-Grad scaling limit and for short times by Lanford, [32]. Such a much more detailed or 'microscopic' version of the second law proves to be valid much beyond the Boltzmann equation; for general theoretical arguments see [26, 27, 28, 18, 20, 21, 52].

From a mathematical point of view, the second law in the form of an H-theorem in fact claims the existence of a Lyapunov functional for a class of evolution equations, and it even hints at how to find that: if we know the underlying microscopic dynamics from which the evolution equation (presumably or provably) follows, one is to search for the Boltzmann entropy. Understanding why this strategy often works brings to the foreground some other important observations: the validity of a macroscopic evolution equation means that there is a *typical* macroscopic behavior in the sense that it is a result of some law of large numbers. The fact that the macroscopic equation is often first order in time means that this macroscopic behavior is *autonomous*. On the other hand, the existence of microscopic configurations violating that typical macroscopic law is not only allowed but in a sense it is even *necessary* for a true irreversible behavior and a strict increase of entropy to occur! When formulated somewhat more precisely, these observations answer various apparent paradoxes as formulated by Loschmidt and Zermelo; a qualitative discussion can be found on various places, see e.g. [33, 6, 26].

Putting these arguments on a mathematically more precise level is relatively simple but it remains very instructive. First, in Section 10, we study a model introduced by Mark Kac, [29]. The arguments are formulated in a substantially more generality in Section 11 (the case of infinite dynamical systems) and Section 12 (the case of large but finite systems). In particular, we explain how an H-theorem follows from the very existence of an (even weakly) autonomous macroscopic dynamics. The resolution of Zermelo's and Loschmidt's paradoxes is understood via a fluctuation symmetry, in this way drawing a link to the first part of these lectures. Most of the presented material and some more details can be found in [11] and references therein.

Since the fundamental laws of nature are presumably quantum, one can further ask how the classical arguments leading to H-theorems need to be changed when starting from a quantum microscopic dynamics. There is no crucial difference up to one important point, namely that the very notion of a macroscopic state needs to be reconsidered because it can be (and in nonequilibrium practice it often *is*) specified through values of mutually incompatible observables. The noncommutativity is a genuine quantum-mechanical feature which cannot be simply waived away by arguments identifying the classical limit with the thermodynamic limit. Furthermore, entropic arguments and microscopic derivations all together play on the level of fluctuations, i.e. before the thermodynamic limit. Starting from Section 13, we explain a possible approach to the quantum problem along the lines of reference [12], thereby generalizing some ideas of John von Neumann, [50]. An interesting side problem is to show how our construction of the quantum Boltzmann entropy relates to other, mathematically simpler but physically *a priori* less plausible constructions. In fact, we show when two definitions of quantum entropies become equivalent in the large system limit. That issue is obviously very related to the problem of quantum large deviations and we will briefly describe the connections. Finally, as an example, we come back to the Kac ring model and we discuss its quantum extension along the lines of reference [10]; see Section 14.

10 Kac Ring Model

There is a simple paradigmatic model introduced by Mark Kac [29] to simplify the mathematics of the Boltzmann equation. While the Boltzmann equation is much more complicated, the Kac model is mathematically simple and free of those extra technical problems that are not really important for understanding some crucial aspects of the emergence of macroscopic irreversibility.

10.1 Microscopic Dynamics

Consider the set $\Lambda = \{1, \ldots, N\}$. We imagine it as a ring in which we identify the sites $1 \equiv N + 1$ and on each site we have one particle and one scatterer. The particles carry a 'spin' $\eta(i) = \pm 1$ and the scatterers can be off or on, $g(i) \in \{0, 1\}$. The resulting set $K = \{-1, 1\}^\Lambda \times \{0, 1\}^\Lambda$ is the state space of our model. The dynamics is deterministic and given via the transformation U on K,

$$(U(\eta, g))(i) = ([1 - 2g(i-1)]\,\eta(i-1),\, g(i)) \mod N \tag{10.1}$$

which generates the (microscopic) deterministic dynamics such that the configuration (η_t, g_t) at time $t = 0, 1, \ldots$ is

$$\eta_t(i) = \eta_0(i-t)[1 - 2g(i-t)] \ldots [1 - 2g(i-1)] \tag{10.2}$$

and $g_t = g$ keeps constant. There is an obvious interpretation: at every time instance t, each spin $\eta_t(i)$ jumps to its successive site, $i+1$, either flipping its value if a scatterer is present, $g(i) = 1$, or keeping its value if $g(i) = 0$. Sampling the initial configuration (η_0, g) from a measure μ_0 on K, the probability to find (η, g) at time t is

$$\mu_t[(\eta, g)] = \mu_0[(\eta_t, g)] \tag{10.3}$$

That is the present variant of the Liouville equation for mechanical systems. Here also the Shannon entropy $S(\mu) = -\sum_{\eta,g} \mu[(\eta, g)] \log \mu[(\eta, g)]$ is time-invariant, $S(\mu_t) = S(\mu_0)$. There is just no strictly increasing Lyapunov function for this dynamical system; in fact, the dynamics is $2N-$periodic. Nevertheless the model exhibits relaxation to equilibrium; to see that, we need to pass to a macroscopic viewpoint.

10.2 Macroscopic Evolution

There are two natural macroscopic observables, the magnetization m^N and the fraction of on-scatterers ρ^N:

$$m^N(\eta) = \frac{1}{N} \sum_{i=1}^{N} \eta(i)\,, \qquad \rho^N(g) = \frac{1}{N} \sum_{i=1}^{N} g(i) \tag{10.4}$$

The emergent *macroscopic* dynamics will have the form

$$(m_t^N, \rho^N) \mapsto (m_{t+1}^N, \rho^N) = \phi(m_t^N, \rho^N) \tag{10.5}$$

at least for very large N. It would imply that the macroscopic data (m_t^N, ρ_t^N) evolve *autonomously*, irrespectively of any actual microscopic configuration (η_t, g_t) that realize (10.4). A simple heuristics[14] suggests $\phi(m, \rho) = ([1 - 2\rho]m, \rho)$ as a candidate map. Yet, it is easy to imagine microscopic configurations that violate that and the question arises how such a macroscopic behavior can/must be understood.

[14] Think of $N(1 \pm m)$ up (down) spins crossing 'on average' $N\rho$ scatterers every time step, entirely neglecting possible time correlations. Such a hand-waving derivation is often referred to as *Stosszahlansatz* (or repeated randomization, molecular chaos approximation,...).

Introducing the counting probability measures

$$\mathbb{P}^N[(\eta, g)] = 2^{-2N} \tag{10.6}$$

and the notation $a \stackrel{\delta}{=} b$ for $|a - b| \leq \delta$, the desired statement has the form of a law of large numbers:[15]

$$\lim_{N\uparrow\infty} \mathbb{P}^N[m^N(\eta_t) \stackrel{\delta}{=} m_0(1-2\rho)^t \,|\, m^N(\eta_0) = m_0;\, \rho^N(g) = \rho] = 1 \tag{10.7}$$

for all $\delta > 0$. This means there is a set of *typical* microscopic configurations satisfying the macroscopic law with map ϕ; those configurations violating that law make a set of zero limit measure. Such a macroevolution is called *autonomous*; note that (10.7) is equivalent to

$$\lim_{N\uparrow\infty} \mathbb{P}^N[\forall t \leq T:\, m^N(\eta_t) \stackrel{\delta}{=} m_0(1-2\rho)^t \,|\, m^N(\eta_0) = m_0;\, \rho^N(g) = \rho] = 1 \tag{10.8}$$

for all $\delta > 0$ and any finite T.

The relaxation to equilibrium along that typical macroevolution is obvious by inspection but one can also construct an explicit witness which is the *Boltzmann entropy* $s(m, \rho)$ defined as the large deviation rate function for the sequence $(m^N, \rho^N)_{N\uparrow+\infty}$ of observables:

$$\mathbb{P}^N[(m^N(\eta), \rho^N(g))] \simeq (m, \rho)] \simeq e^{Ns(m,\rho)} \tag{10.9}$$

This is to be understood in the logarithmic sense after taking the limit $N \uparrow +\infty$, i.e., it is a shorthand for the limit statement

$$s(m, \rho) = \lim_{\delta\downarrow 0} \lim_{N\uparrow+\infty} \frac{1}{N} \log \mathbb{P}^N[m^N(\eta) \stackrel{\delta}{=} m;\, \rho^N(\eta) \stackrel{\delta}{=} \rho] \tag{10.10}$$

This is simply the binomial entropy,

$$s(m,\rho) = \begin{cases} -\frac{1+m}{2}\log(1+m) - \frac{1-m}{2}\log(1-m) - \rho\log 2\rho \\ \qquad\qquad -(1-\rho)\log 2(1-\rho) \;\text{ if }\; -1 < m < 1,\; 0 < \rho < 1 \\ -\infty \qquad\qquad\qquad\qquad\qquad \text{otherwise} \end{cases} \tag{10.11}$$

and one checks that $s(\phi(m, \rho)) > s(m, \rho)$ whenever $m \neq 0$ (system off equilibrium) and $0 < \rho < 1$ (nonsingular macrodynamics). Using the notation $m_t = m_0(1-2\rho)^t$, it yields that $s(m_t, \rho)$ is a strictly increasing

[15] In fact, a strong law of large numbers is also true for this model, see [29], but the weak law is sufficient for our purposes. Later we will meet an even substantially weaker autonomy condition.

function of time. Following Boltzmann's terminology, such a statement is called an *H-theorem*; the Boltzmann entropy is a Lyapunov function.

It should be clear that we say nothing yet about possible macroevolutions corresponding to those exceptional microstates not verifying the macroscopic map ϕ. That will come in Section 10.3.

Proof of (10.7): Use the shorthand

$$\tilde{\mathbb{E}}^N[\,\cdot\,] = \mathbb{E}^N[\,\cdot\,|\,m^N(\eta_0) = m_0;\,\rho^N(g) = \rho]$$

for the expectations conditioned on the initial macrostate (m_0, ρ). One easily checks that (1) $\mathbb{P}^N$ are permutation-invariant measures, (2) η_0 and g are independently distributed under $\mathbb{P}^N$, and (3) the following asymptotic decoupling is true:[16] provided that $-1 < m_0 < 1$ and $0 < \rho < 1$, there is a sequence $\Delta_N^{(k)}$, $\lim_N \Delta_N^{(k)} = 0$ for all $k = 1, 2, \ldots$, such that

$$|\tilde{\mathbb{E}}^N[\eta(i_1)\ldots\eta(i_k)] - (\tilde{\mathbb{E}}^N[\eta(1)])^k| \le \Delta_N^{(k)} \tag{10.12}$$

and

$$|\tilde{\mathbb{E}}^N[g(i_1)\ldots g(i_k)] - (\tilde{\mathbb{E}}^N[g(1)])^k| \le \Delta_N^{(k)} \tag{10.13}$$

for all $1 \le i_1 < i_2 < \ldots < i_k \le N$.

Then, one subsequently gets

$$\begin{aligned}\tilde{\mathbb{E}}^N[m^N(\eta_t)] &= \frac{1}{N}\sum_{i=1}^N \tilde{\mathbb{E}}^N[\eta_0(i-t)\{1-2g(i-t)\}\ldots\{1-2g(i-1)\}]\\ &= \tilde{\mathbb{E}}^N[\eta_0(1)]\,\tilde{\mathbb{E}}^N[\{1-2g(1)\}\ldots\{1-2g(t)\}]\\ &= m_0\{(1-2\,\tilde{\mathbb{E}}^N[g(1)])^t + A_N\}\\ &= m_0(1-2\rho)^t + o(1)\end{aligned} \tag{10.14}$$

by using that $|A_N| \le \Delta_N^{(t)}$ due to (10.13). Similarly for the second moment,

$$\begin{aligned}\tilde{\mathbb{E}}^N[(m^N(\eta_t))^2] = \frac{1}{N}\sum_{i=1}^N \tilde{\mathbb{E}}^N[\eta_0(1)\,\eta_0(i)]\,\tilde{\mathbb{E}}^N[\{1-2g(1)\}\ldots\{1-2g(t)\}\\ \times\{1-2g(i)\}\ldots\{1-2g(i+t-1)\}]\end{aligned} \tag{10.15}$$

[16] This property can be recognized as an instance of the equivalence between microcanonical and canonical ensembles.

Observe that for $t+1 \le i \le N+1-t$ there is no pair of the g's in the above product, acting on the same site. Hence, using the permutation-invariance and the asymptotic decoupling (10.12)–(10.13),

$$\begin{aligned}\tilde{\mathbb{E}}^N[(m^N(\eta_t))^2] &= \frac{1}{N}\sum_{i=t+1}^{N+1-t}\{m_0^2 + B_N(i)\}\{(1-2\,\tilde{\mathbb{E}}^N[g(1)])^{2t} + C_N(i)\} \\ &\quad + \frac{1}{N}\Big(\sum_{i=1}^{t} + \sum_{i=N+2-t}^{N}\Big)\{m_0^2 + B_N(i)\} \\ &\quad \times \tilde{\mathbb{E}}^N[\{1-2g(1)\}\dots\{1-2g(t)\} \\ &\qquad \times \{1-2g(i)\}\dots\{1-2g(i+t-1)\}] \\ &= \frac{N-2t+1}{N} m_0^2(1-2\rho)^{2t} + o(1) \\ &= m_0^2(1-2\rho)^{2t} + o(1) \end{aligned} \tag{10.16}$$

since the remainders satisfy $|B_N(i)| \le \Delta_N^{(2t)}$, $|C_N(i)| \le \Delta_N^{(2t)}$ for all i, and using a simple bound on the last term. The weak law of large numbers (10.7) then follows from (10.14) and (10.16) via a Chebyshev inequality.

Remark that there is a considerable freedom in the choice of the measures from which the initial configurations are sampled. The 'microcanonical' measure

$$\tilde{\mathbb{P}}^N[\,\cdot\,] = \mathbb{P}^N[\,\cdot\,|\,m^N(\eta) = m_0;\, \rho^N(g) = \rho] \tag{10.17}$$

is most natural but for obtaining (10.7), it can be replaced by various other ensembles. There is for example the 'canonical' measure

$$\mathbb{P}^N_{\mathrm{can}}[(\eta, g)] = \frac{1}{\mathcal{Z}^N} \exp \sum_{i=1}^{N} (\beta\eta(i) + \alpha g(i)) \tag{10.18}$$

with the Lagrange multipliers β, α being fixed by the conditions

$$\mathbb{E}^N_{\mathrm{can}}[\eta(1)] = m_0\,, \qquad \mathbb{E}^N_{\mathrm{can}}[g(1)] = \rho$$

and $\mathcal{Z}^N$ is the normalization factor. It is easy to check that the autonomy (10.7) remains true if replacing $\tilde{\mathbb{P}}^N$ with $\mathbb{P}^N_{\mathrm{can}}$; in fact, the proof is simpler now since $\mathbb{P}^N_{\mathrm{can}}$ exactly factorizes and hence the above remainders A_N, B_N, C_N as in (10.14) and (10.16) are zero. The corresponding large deviation rate function $s_{\mathrm{can}}(\eta, g)$ which enters the law

$$\mathbb{P}^N_{\rm can}[(m^N(\eta),\rho^N(g)) \simeq (m,\rho)] \simeq e^{Ns_{\rm can}(m,\rho)} \tag{10.19}$$

can also be computed (the easiest via the Gartner-Ellis theorem, [9]) with the result $s_{\rm can}(m,\rho) = s(m,\rho)$. This equality shows the equivalence of ensembles on the level of entropies, which is well studied in equilibrium statistical physics. We come back to the problem of equivalence within a quantum framework and in a substantially larger generality in Section 13.4.

10.3 Irreversibility and Entropy Production

Consider a modification of the microdynamics (10.1) in which the particles jump to the left instead of to the right. It is given by the map

$$(\bar U(\eta,g))(i) = ([1-2g(i)]\,\eta(i+1),\, g(i)) \mod N \tag{10.20}$$

which is an inverse of U, i.e., $\bar U \circ U = U \circ \bar U = 1$. The spin configuration at time t as evolved from η through the dynamics $\bar U$ is denoted by $\bar\eta_t$:

$$\bar\eta_t(i) = \eta_0(i+t)[1-2g(i+t-1)]\ldots[1-2g(i)] \tag{10.21}$$

Observe that the sequence (trajectory) $(\eta_0,\eta_1,\ldots,\eta_t)$ is allowed (possible) under the original microscopic dynamics iff $(\eta_t,\eta_{t-1},\ldots,\eta_0)$ is possible under U. That invertibility is referred to as *dynamical (time-) reversibility*. It can be formulated differently by extending the configuration space K with a 'velocity' variable $v \in \{-1,1\}$ and by defining the dynamics via the transformation

$$V(\eta,g,v) = \begin{cases} (U(\eta,g),v) & \text{if } v=+1 \\ (\bar U(\eta,g),v) & \text{if } v=-1 \end{cases} \tag{10.22}$$

With the involution π, $\pi(\eta,g,v) = (\eta,g,-v)$ the dynamical reversibility gets the form: $\pi\circ V\circ\pi = V^{-1}$; the time-reversed microscopic dynamics is then achieved by inverting the velocity v.

The macroscopic time-evolution ϕ in which $m \mapsto (1-2\rho)m \in [-1,1]$ is invertible as well (provided that $\rho \neq \frac12$). Yet, the typical macroscopic time-evolution corresponding to $\bar U$ is not ϕ^{-1} but rather ϕ again, i.e., the law of large numbers (10.7) stays true when replacing U and η_t with $\bar U$ and $\bar\eta_t$! It simply means that the macroscopic evolution $m \mapsto m(1-2\rho)$ does *not* get inverted by starting from a *typical* microscopic configuration η corresponding to the macroscopic state $m(1-2\rho)$ and by applying the inverted microscopic dynamics (or by inverting the velocity). Naturally, there exist microscopic configurations η for which

the inverted macroevolution $m(1-2\rho) \mapsto m$ along the dynamics $\bar{U}$ would be observed—this is precisely what the dynamical reversibility claims—they show up to be exceedingly *exceptional* under macrostate $m(1-2\rho)$, however. This physical impossibility to invert the macroscopic evolution is referred to as *macroscopic irreversibility*.

The macroscopic irreversibility in the above sense on the one hand and the strict increase of the Boltzmann entropy on the other hand, are often used as synonyms. Let us formulate their relation a bit more precisely. Observe that the two sets

$$\{(\eta_0, g) : m^N(\eta_0) \stackrel{\delta}{=} m_0;\, m^N(\eta_t) \stackrel{\delta}{=} m_t;\, \rho^N(g) \stackrel{\delta}{=} \rho\}$$

and

$$\{(\eta_0, g) : m^N(\bar{\eta}_t) \stackrel{\delta}{=} m_0;\, m^N(\eta_0) \stackrel{\delta}{=} m_t;\, \rho^N(g) \stackrel{\delta}{=} \rho\}$$

with $m_t = m_0(1-2\rho)^t$ have the same cardinalities (check that the map $\bar{U}$ is a bijection between these sets). Hence, they have the same measures under $\mathbb{P}^N$, which we write as

$$\begin{aligned} &\log \mathbb{P}^N[m^N(\eta_t) \stackrel{\delta}{=} m_t \,|\, m^N(\eta_0) \stackrel{\delta}{=} m_0;\, \rho^N(g) \stackrel{\delta}{=} \rho] \\ &\qquad + \log \mathbb{P}^N[m^N(\eta_0) \stackrel{\delta}{=} m_0;\, \rho^N(g) \stackrel{\delta}{=} \rho] \\ &= \log \mathbb{P}^N[m^N(\bar{\eta}_t) \stackrel{\delta}{=} m_0 \,|\, m^N(\eta_0) \stackrel{\delta}{=} m_t;\, \rho^N(g) \stackrel{\delta}{=} \rho] \\ &\qquad + \log \mathbb{P}^N[m^N(\eta_0) \stackrel{\delta}{=} m_t;\, \rho^N(g) \stackrel{\delta}{=} \rho] \end{aligned} \tag{10.23}$$

Dividing by N, taking the limits $N \uparrow \infty$ and $\delta \downarrow 0$ in this order, and using the law of large numbers (10.7), we get the large deviation law

$$\begin{aligned} &\lim_{\delta \downarrow 0} \lim_{N \uparrow \infty} \frac{1}{N} \log \mathbb{P}^N[m^N(\bar{\eta}_t) \stackrel{\delta}{=} m_0 \,|\, m^N(\eta_0) \stackrel{\delta}{=} m_t;\, \rho^N(g) \stackrel{\delta}{=} \rho] \\ &\qquad = s(m_0, \rho) - s(m_t, \rho) \end{aligned} \tag{10.24}$$

or

$$\mathbb{P}^N[m^N(\bar{\eta}_t) \simeq m_0 \,|\, m^N(\eta_0) \simeq m_t;\, \rho^N(g) \simeq \rho] \simeq e^{-N[s(m_t,\rho) - s(m_0,\rho)]} \tag{10.25}$$

which is quite a remarkable relation. Notice first that it provides another derivation of the H-theorem: since the left-hand side is less than one, one immediately gets $s(m_t, \rho) \geq s(m_0, \rho)$. Further, the left-hand side is nothing but the probability that a configuration η_0 sampled from macrostate m_t and evolved according to $\bar{U}$ (or equivalently U for this model), exhibits the macroscopic transition from m_t to m_0, which is just an inversion of the typical transition $m_0 \mapsto m_t$.

Macroscopic irreversibility amounts to the statement that such inverted macroscopic transitions are physically impossible; here we have a quantitative evaluation how rare they really are: the large deviation rate function for the backward transition $m_t \mapsto m_0$ with respect to $\bar{U}$ just coincides with the entropy production along the *typical* macroevolution $m_0 \mapsto m_t$ with respect to microscopic dynamics U. Inverting the logic, this can be read off as a formula for the entropy production, possibly useful provided that the probability of those rare backward transitions can be evaluated or estimated.

The seeming inconsistency between the microscopic reversibility and the macroscopic irreversibility is known as the Loschmidt paradox. Equality (10.25) in a sense solves this paradox and put it in a correct perspective: those macroscopic trajectories obtained by time-reverting the typical ones $(\phi^t(m);\ t = 0, 1, \ldots)$ are indeed observable for finite N, however they are exponentially damped. Notice that (10.25) is nothing but a macroscopic analogue of the detailed balance condition (4.4) that we have already discussed in the context of lattice gases.

11 Infinite Systems

The above analysis of the Kac model shows up to be quite generic and one can easily extend those arguments to a more general setup. The aim of the present section is to formulate general sufficient conditions for the existence of a Lyapunov function for a class of macroscopic dynamics, or, equivalently, for an H-theorem to be valid.

11.1 Dynamical Systems, Macrostates, and Entropy

On a microscopic level, we consider a family of classical dynamical systems $(K^N, U_t^N, \mathbb{P}^N)_{N\uparrow+\infty}$, where the label N should be thought of as a spatial extension of the system and the maps[17] $(U_t^N)_{t\geq 0}$ are assumed to satisfy the semigroup condition $U_t^N U_s^N = U_{t+s}$ for all $t, s \geq 0$. The probability measures $\mathbb{P}^N$ are invariant under the dynamics: $\mathbb{P}^N (U^N)^{-1} = \mathbb{P}^N$.

The macroscopic level of description is specified by a collection of macroscopic observables. These are some maps $M^N : K^N \mapsto \Omega$ into a metric space (Ω, d) of *macrostates*. We assign to every $m \in \Omega$ the

[17] Depending on an application, the maps U_t^N can be e.g. Hamiltonian flows, possibly with the time and space suitably rescaled with N. The details are not really important for what follows, we will only require the microscopic dynamics to satisfy a few general conditions, see below.

Boltzmann entropy defined as the large deviation rate function under the measures $\mathbb{P}^N$:

$$s(m) = \lim_{\delta \downarrow 0} \limsup_{N \uparrow +\infty} \frac{1}{N} \log \mathbb{P}^N [M^N(x) \stackrel{\delta}{=} m] \,, \qquad m \in \Omega \tag{11.1}$$

using the shorthand $m \stackrel{\delta}{=} m'$ whenever $d(m, m') \leq \delta$. Denote

$$\Omega_0 = \{m \in \Omega;\, s(m) > -\infty\} \tag{11.2}$$

the set of those macrostates that are admissible; we assume $\Omega_0 \neq \emptyset$. Note this is a first nontrivial assumption: the parametrization by N and an eventual rescaling have to be meaningful so that (M^N) indeed satisfies the large deviation principle with a finite rate function $s(m)$ on some large enough space Ω_0.

11.2 Autonomous Evolution and H-theorem

Starting from a microscopic configuration $x \in K^N$, the macroscopic trajectory is simply the collection $(M^N(U_t^N x))_{t \geq 0}$. We assume the existence of an *autonomous* macroscopic dynamics in the following sense: let there be a collection $(\phi_t)_{t \geq 0}$ of maps $\phi : \Omega_0 \mapsto \Omega_0$ satisfying

1. the semigroup condition

$$\phi_t \circ \phi_s = \phi_{t+s} \,, \qquad t, s \geq 0 \tag{11.3}$$

2. a weak autonomy condition

$$\lim_{\delta \downarrow 0} \lim_{N \uparrow +\infty} \frac{1}{N} \log \mathbb{P}^N [M^N(U_t^N x) \stackrel{\delta}{=} \phi_t(m) \,|\, M^N(x) \stackrel{\delta}{=} m] = 0 \tag{11.4}$$

for all $t \geq 0$ and $m \in \Omega_0$.

Notice that (11.4) is a much weaker condition than the law of large numbers (10.7) valid for the Kac model. In particular, no typical macroscopic evolution is required to exist; that $(\phi_t)_{t \geq 0}$ can e.g. be a single realization of a stochastic process describing a macroscopic evolution of a system passing through a number of branching points.[18] On the other hand, this condition is generally not satisfied by stochastic systems on mesoscopic scale and/or without involving the large N limit. In that

[18] To have in mind a specific scenario, think of a ferromagnet being cooled down from a high-temperature paramagnetic state. When passing the critical temperature, the system randomly (= depending on the initial microscopic configuration) chooses one of the ferromagnetic states with broken symmetry.

sense, condition (11.4) draws a sharp border line between macroscopic and mesoscopic systems.

Since $\mathbb{P}^N$ is invariant under U_t^N, conditions (11.3)–(11.4) are equivalent with a single condition

$$\lim_{\delta\downarrow 0}\lim_{N\uparrow+\infty}\frac{1}{N}\log\mathbb{P}^N[M^N(U_t^N x)\stackrel{\delta}{=}\phi_t(m)\,|\,M^N(U_s x)\stackrel{\delta}{=}\phi_s(m)]=0 \tag{11.5}$$

required for all $t\geq s\geq 0$. Using the invariance $\mathbb{P}^N$ again, we find that for any pair $m,m'\in\Omega_0$ of macrostates,

$$\begin{aligned}\log\mathbb{P}^N(M^N(x)\stackrel{\delta}{=}m') &= \log\mathbb{P}^N(M^N(U_t^N x)\stackrel{\delta}{=}m')\\ &\geq \log\mathbb{P}^N(M^N(U_t^N x)\stackrel{\delta}{=}m'\,|\,M^N(U_s^N x)\stackrel{\delta}{=}m)\\ &\quad+\log\mathbb{P}^N(M^N(U_s^N x)\stackrel{\delta}{=}m)\end{aligned} \tag{11.6}$$

which we are again going to divide by N, to take the upper limit $N\uparrow+\infty$ and then to take the limit $\delta\downarrow 0$. Choosing first $m=\phi_s(m)$ and $m'=\phi_t(m)$, autonomy condition (11.5) yields

$$s(\phi_t(m))\geq s(\phi_s(m))\,,\qquad t\geq s\geq 0 \tag{11.7}$$

which is an H-theorem. Second, for $m=\phi_t(m)$ and $m'=\phi_s(m)$ it yields the inequality

$$\begin{aligned}-\lim_{\delta\downarrow 0}\limsup_{N\uparrow+\infty}\frac{1}{N}\log\mathbb{P}^N(M^N(U_t^N x)\stackrel{\delta}{=}\phi_s(m)\,|\,M^N(U_s^N x)\stackrel{\delta}{=}\phi_t(m))&\\ \geq s(\phi_t(m))-s(\phi_s(m))&\end{aligned} \tag{11.8}$$

again for all $t\geq s\geq 0$, which provides an upper bound on the Boltzmann entropy production.

For an invertible microdynamics (U_t^N) inequalities (11.7)–(11.8) can be turned into a single equality by essentially repeating the computation of Section 10.3, see (10.23)–(10.24). The result reads

$$s(\phi_t(m))-s(\phi_s(m))=\bar{J}_{s,t}(\phi_t(m),\phi_s(m))\geq 0 \tag{11.9}$$

where

$$\begin{aligned}&-\bar{J}_{s,t}(m,m')\\ &=\lim_{\delta\downarrow 0}\limsup_{N\uparrow+\infty}\frac{1}{N}\log\mathbb{P}^N[M^N(\bar{U}_t^N x)\stackrel{\delta}{=}m'\,|\,M^N(\bar{U}_s^N x)\stackrel{\delta}{=}m]\end{aligned} \tag{11.10}$$

is the rate function for the transition from macrostate m at time s to macrostate m' at time t along the time-reversed dynamics $\bar{U}_t^N \equiv (U_t^N)^{-1}$. The conclusions of Section 10.3 apply as well in this general case: $\bar{J}_{s,t}(\phi_t(m), \phi_s(m))$ is a natural 'measure' of macroscopic irreversibility, and we have proven it is just equal to the entropy production for the transition $\phi_s(m) \mapsto \phi_t(m)$ fulfilling the autonomy (11.5).

Semigroup condition (11.3) is crucial and cannot be simply relaxed. Indeed, assuming only the autonomy in the form (11.4), one would still have the inequality between the initial and final Boltzmann entropies: $s(\phi_t(m)) \geq s(m)$, $t \geq 0$, however, $s(m_t)$ might not be monotone in general. As an example, think of the macrodynamics $\phi_t : \mathbb{R} \mapsto \mathbb{R}$ given as $\phi_t(m) = m\, r^t \cos \omega t$, $|r| < 1$, which is like the position of an underdamped pendulum swinging around its equilibrium position.

The missing semigroup property can be recovered by including additional macroscopic observables; in the case of the pendulum one would naturally add its velocity as another observable. To conclude, the semigroup condition is basically a restriction on the choice of the collection of macroscopic observables, which needs to be in that sense 'complete.'

12 Finite Systems

In this section we evaluate the necessity of the large N limit in the above arguments, and we attempt a microscopic formulation of the H-theorem for an entropy defined upon a finite system and as a functional on microscopic configurations.

12.1 Zermelo-Poincaré Paradox

For any fixed N the dynamical system $(K, U_t, \mathbb{P}) \equiv (K^N, U_t^N, \mathbb{P}^N)$ is really a *finite size* system; this is encoded by the assumption that $\mathbb{P}$ is a probability (and hence normalizable) measure. For simplicity, we consider in this section a macroscopic observable $M : K \mapsto \Omega$ such that

$$\Omega_0 = \{m \in \Omega;\, \mathbb{P}[M^{-1}(m)] > 0\} \tag{12.1}$$

is finite or countable. As entropy function we take

$$S(m) = \log \mathbb{P}[M^{-1}(m)]\,, \qquad m \in \Omega_0 \tag{12.2}$$

The well-known Poincaré recurrence theorem then reads that for $\mathbb{P}-$almost every microstate $x \in K$ and any time t_0 one has $M(U_t x) = M(x)$ for some $t > t_0$, i.e., the trajectory almost surely returns back to

the initial macrostate $M(x)$. This is usually phrased as the impossibility for the entropy (12.10) to be an increasing (nonconstant) function of time, which is also known as the Zermelo(-Poincaré) paradox. However, the analysis in the previous section and the Kac example give a clue: the recurrence time increases with the system size N and is shifted away to infinity in the thermodynamic limit.[19]

In the context of H-theorems it is instructive to reformulate Zermelo's objection in still a slightly modified way. There is in fact a tempting 'trivialization' of the argument leading to the H-theorem which goes as follows:

For our finite system, the autonomy could simply mean that the macroscopic evolution as specified by a map $\phi : \Omega_0 \mapsto \Omega_0$ is just what takes place for almost every microstate, i.e., that for $\mathbb{P}-$a.e. $x \in K$ one has

$$M(U_t x) = \phi_t(M(x)) \,, \qquad t \geq 0 \tag{12.3}$$

Were this indeed true as such, it would automatically imply the semigroup condition since, almost surely,

$$M(U_{t+s}x) = \phi_t(M(U_s x)) = \phi_t \circ \phi_s(M(x)) \,, \qquad t, s \geq 0 \tag{12.4}$$

Second, it would mean that

$$\mathbb{P}[(U_t)^{-1} M^{-1}(m_t) \cap M^{-1}(m)] = \mathbb{P}[M^{-1}(m)] \tag{12.5}$$

with $m_t = \phi_t(m)$, i.e., the set of microstates $M^{-1}(m)$ evolves to a subset of $M^{-1}(m_t)$, up to a zero measure set. Hence,

$$S(m_t) = \log \mathbb{P}[(U_t)^{-1} M^{-1}(m_t)] \geq \mathbb{P}[M^{-1}(m)] = S(m) \tag{12.6}$$

due to the invariance of $\mathbb{P}$. Finally,

$$S(m_t) = S(\phi_{t-s}(m_s)) \geq S(m_s) \,, \qquad t \geq s \geq 0 \tag{12.7}$$

On the other hand, by the above Poincaré recurrence, $S(m_t) = S(m)$ for infinitely many t, and hence $S(m_t)$ is constant!

The above computation shows that the assumption of autonomy in the form (12.3) or (12.5) is too strong. If fulfilled, the macroscopic evolution would necessarily be reversible and the entropy constant. Our condition of autonomy (11.4) is much weaker and it does not guarantee the semigroup property. In the Kac example, the law of large numbers (10.7) is far stronger than autonomy (11.4), yet still consistent with a macroscopic irreversible evolution, as we have checked explicitly.

[19] For the Kac example the recurrence time is of order N. However, in many dynamical systems it goes exponentially with the number of degrees of freedom, which is itself $e^{O(N)}$.

12.2 Microscopic H-theorem

We come back to the general framework of Section 11.1 and consider again a sequence of dynamical systems $(K^N, U_t^N, \mathbb{P}^N)_N$ and a general macroscopic observable $M^N : K^N \to \Omega$, macroscopic in the sense that Ω_0 defined in (11.1) is nonempty. Our aim is now to formulate an H-theorem for Boltzmann entropy assigned to each *microstate* of a single, possibly large finite-size system. Put it differently, we want to see how much the entropy is allowed to fluctuate around a monotone path when evaluated along a single microscopic trajectory of a single dynamical system with N large but fixed.

For N fixed the Boltzmann entropy is no longer unambiguously defined since the sets $\{x;\, M^N(x) \stackrel{\delta}{=} m\}$ depend on the width $\delta > 0$ and are not necessarily all disjoint for different macroscopic states m. To be safe we assign to every microstate $x \in K^N$ the interval $[S_<^{N,\delta}(x), S_>^{N,\delta}(x)]$ of entropies defined as

$$S_<^{N,\delta}(x) = \inf_{m\in\Omega_0} \{S^{N,\delta}(m);\, m \stackrel{\delta}{=} M^N(x)\} \tag{12.8}$$

$$S_>^{N,\delta}(x) = \sup_{m\in\Omega_0} \{S^{N,\delta}(m);\, m \stackrel{\delta}{=} M^N(x)\} \tag{12.9}$$

where

$$S^{N,\delta}(m) = \log \mathbb{P}^N[M^N(x) \stackrel{\delta}{=} m] \tag{12.10}$$

We impose the autonomy assumption in the form of a law of large numbers:

$$\lim_{\delta\downarrow 0} \lim_{N\uparrow+\infty} \mathbb{P}^N[M^N(U_t^N x) \stackrel{\delta}{=} \phi_t(m) \,|\, M^N(x) \stackrel{\delta}{=} m] = 1 \tag{12.11}$$

for all $m \in \Omega_0$, $t \geq 0$, and with maps $\phi_t : \Omega_0 \mapsto \Omega_0$ such that $\phi_t \circ \phi_s = \phi_{t+s}$, $t \geq s \geq 0$.

Let us fix some initial condition $m \in \Omega_0$, $\delta > 0$ and a finite sequence of times $0 = t_0 < t_1 < \ldots < t_Q$. Combining assumption (12.11) with the semigroup condition, the remainder

$$D^{N,\delta}(s,t;m) := 1 - \mathbb{P}^N[M^N(U_t^N x) \stackrel{\delta}{=} \phi_t(m) \,|\, M^N(U_s^N x) \stackrel{\delta}{=} \phi_s(m)] \tag{12.12}$$

satisfies $\lim_{\delta\downarrow 0} \lim_{N\uparrow+\infty} D^{N,\delta}(s,t;m) = 0$ whenever $t \geq s \geq 0$. By subadditivity,

$$\mathbb{P}^N[M^N(U^N_{t_j}x) \stackrel{\delta}{=} \phi_{t_j}(m),\, j=1,\ldots,Q \,|\, M^N(x) \stackrel{\delta}{=} m] \geq 1 - \sum_{j=1}^{Q} D^{N,\delta}(0,t_j;m) \quad (12.13)$$

Whenever $m^N(U_t x) \stackrel{\delta}{=} \phi_t(m)$ then

$$S^{N,\delta}_<(U^N_t x) \leq S^{N,\delta}(\phi_t(m)) \leq S^{N,\delta}_>(U^N_t x) \quad (12.14)$$

As a consequence, (12.13) gives

$$\mathbb{P}^N[S^{N,\delta}_<(U_{t_j}x) \leq S^{N,\delta}(\phi_{t_j}(m)) \leq S^{N,\delta}_>(U_{t_j}x), j=1,\ldots,Q \,|\, M^N(x) \stackrel{\delta}{=} m] \geq 1 - \sum_{j=1}^{Q} D^{N,\delta}(0,t_j;m) \quad (12.15)$$

Entropies at successive times satisfy the inequality following from (12.12):

$$S^{N,\delta}(\phi_{t_j}(m)) \geq S^{N,\delta}(\phi_{t_{j-1}}(m)) + \log(1 - D^{N,\delta}(t_{j-1},t_j;m)) \quad (12.16)$$

Using that

$$\lim_{\delta\downarrow 0} \lim_{N\uparrow+\infty} \min_{j=1}^{Q} \log(1 - D^{N,\delta}(t_{j-1},t_j;m)) = 0 \quad (12.17)$$

inequalities (12.15)–(12.16) yield the main result of this section:

For any $\Delta > 0$, $m \in \Omega_0$, and a finite sequence $0 = t_0 < t_1 < \ldots t_Q$ of times, one has

$$\lim_{\delta\downarrow 0} \lim_{N\uparrow+\infty} \mathbb{P}^N[S^{N,\delta}_>(U_{t_j}x) \geq S^{N,\delta}_<(U_{t_{j-1}}x) - \Delta, j=1,\ldots,Q | M^N(x) \stackrel{\delta}{=} m] = 1 \quad (12.18)$$

Therefore, the finite-system entropy violates the monotonicity as little as required with probability arbitrarily closed to one, provided that δ is small enough and N large enough. This is the announced microscopic H-theorem.

Remark that the ambiguity with the definition of finite-system entropy does not arise if the observables M^N take only finitely or countably many values (as in Section 12.1 but this time uniformly for all N). In that case, one can set $\delta = 0$ and the entropy is simply

$$S^N(x) := S^{N,0}_<(x) = S^{N,0}_>(x) = \log \mathbb{P}^N[(M^N)^{-1}M^N(x)] \quad (12.19)$$

(Compare with (12.2).) The microscopic H-theorem (12.18) then becomes, under the same assumptions,

$$\lim_{N\uparrow+\infty} \mathbb{P}^N[S^N(U_{t_j}x) \geq S^N(U_{t_{j-1}}x) - \Delta,\, j = 1,\ldots,Q \,|\, M^N(x) = m] = 1 \tag{12.20}$$

13 Quantum Systems

The arguments presented in the previous section do not dramatically change when passing from a classical to a quantum dynamics. What does need to be refined however, is the very description of macroscopic states due to the inherent incompatibility of quantum observables, as visible from the noncommutativity of corresponding (self-adjoint) operators before any macroscopic limit is taken.

Obviously, the question of a quantum fluctuation theory and of quantum limiting behavior is not restricted to nonequilibrium physics. The difference between equilibrium and nonequilibrium macroscopic states lies mainly in the choice of macroscopic constraints. The constraints describing equilibrium (like energy, particle number) usually commute and hence the problem we discuss here typically falls in a nonequilibrium context.

A first attack on this problem dates back to John von Neumann, [50]. His idea went as follows: The single particle position and momentum operators Q, P satisfy the commutation relation[20] $[Q,P] = i$. Hence, assigning copies Q_i, P_i, $i = 1,\ldots,N$ to each of N particles, the averages $Q^N = \frac{1}{N}\sum_i Q_i$, $P^N = \frac{1}{N}\sum_i P_i$ satisfy $[Q^N, P^N] = \frac{i}{N}$. Although they do not commute and hence cannot be diagonalized together (nor simultaneously measured), one can think of suitable modifications $\tilde{Q}^N$, $\tilde{P}^N$ that already commute and that in a sense well approximate the originals, at least for large N. Indeed, von Neumann explicitly constructs commuting operators $\tilde{Q}^N$, $\tilde{P}^N$ which have purely discrete spectra of nondegenerate eigenvalues $(q_\alpha^N, p_\alpha^N)_\alpha$, and whose eigenvectors $(\psi_\alpha^N)_\alpha$ make a complete orthonormal basis system in $L^2(\mathbb{R}^N)$. They approximate the operators Q^N, P^N in the sense that

$$(\psi_\alpha^N, Q^N\psi_\alpha^N) = q_\alpha^N\,, \qquad (\psi_\alpha^N, P^N\psi_\alpha^N) = p_\alpha^N \tag{13.1}$$

and

$$\|(Q^N - q_\alpha^N)\,\psi_\alpha^N\| \leq \frac{C}{\sqrt{N}}\,, \qquad \|(P^N - p_\alpha^N)\,\psi_\alpha^N\| \leq \frac{C}{\sqrt{N}} \tag{13.2}$$

with $C \approx 60$, see [50] for details.

[20] Set the Planck constant to one.

What follows is a slight modification and generalization of the above idea that comes close to the point of view of quantum information theory. Instead of modifying the operators themselves, we look for the largest or typical subspaces that in some sense well approximate the eigenspaces for given eigenvalues, simultaneously for *all* macroscopic observables from a collection. This construction proves to be natural since the corresponding entropy, measuring the dimension of that typical subspace, actually satisfies a variational principle. Hence, it can be directly compared with another, more familiar although physically less satisfactory construction based on maximizing the von Neumann entropy. The (non)equality of both entropies is then a problem of (non)equivalence of ensembles. Looked at from another angle, such an equivalence gives a counting interpretation to the von Neumann entropy in the thermodynamic limit, and opens interesting possibilities towards a consistent and meaningful scheme of quantum large deviations.

13.1 Quantum Macrostates and Entropy

A macroscopically large quantum system is modeled by a sequence of finite-dimensional Hilbert spaces $(\mathcal{H}^N)_{N\uparrow+\infty}$ on which we have standard traces Tr^N. As macroscopic observables we consider a collection $M^N = (M_k^N)_{k\in I}$ of self-adjoint operators on $\mathcal{H}^N$; for simplicity, we assume I to be finite. For each operator there is a projection-valued measure $\mathcal{Q}_k^N$ on $\mathbb{R}$ such that, by the spectral theorem,

$$F(M_k^N) = \int_{\mathbb{R}} \mathcal{Q}_k^N(\mathrm{d}z)\, F(z)\,, \qquad F \in C(\mathbb{R}) \tag{13.3}$$

(which is just to say that M_k^N is unitarily equivalent to a multiplication, or simply that M_k^N can be diagonalized.) A quantum counterpart of the classical set of microstates $M_k^N \stackrel{\delta}{=} m_k$ for some macrostate $m_k \in \mathbb{R}$ is the projection

$$\mathcal{Q}_k^{N,\delta}(m_k) = \int_{m_k-\delta}^{m_k+\delta} \mathcal{Q}_k^N(\mathrm{d}z) \tag{13.4}$$

Commuting Observables

As a warm-up, assume first that M^N is a collection of mutually commuting operators. In that case, $\mathcal{Q}^N(\mathrm{d}z) = \prod_{k\in I} \mathcal{Q}_k^N(\mathrm{d}z_k)$ is a common projection-valued measure, (13.3) extends to

$$F(M^N) = \int_{\mathbb{R}^I} \mathcal{Q}^N(\mathrm{d}z)\, F(z)\,, \qquad F \in C(\mathbb{R}^I) \tag{13.5}$$

and a macrostate $m = (m_k)_{k\in I}$ gets represented by the projection

$$\mathcal{Q}^{N,\delta}(m) = \prod_{k\in I} \mathcal{Q}_k^{N,\delta}(m_k) \tag{13.6}$$

The classical entropy, say in the form (12.10), extends to

$$S^{N,\delta}(m) = \log \mathrm{Tr}^N[\mathcal{Q}^{N,\delta}(m)] \tag{13.7}$$

This is a formalism entirely equivalent to the one for classical systems.

General Observables

For a general collection M^N of observables and a macrostate $m \in \mathbb{R}^I$, projections $(\mathcal{P}^N)_{N\uparrow+\infty}$ are said to be *concentrating* at m, written $\mathcal{P}^N \to m$, whenever

$$\lim_{N\uparrow+\infty} \mathrm{tr}^N[F(M_k^N)\,|\,\mathcal{P}^N] = F(m_k) \tag{13.8}$$

is satisfied for all $F \in C(\mathbb{R})$ and $k \in I$, with the notation

$$\mathrm{tr}^N[\,\cdot\,|\,\mathcal{P}^N] = \frac{\mathrm{Tr}^N[\mathcal{P}^N\cdot\,]}{\mathrm{Tr}^N[\mathcal{P}^N]} \tag{13.9}$$

for the normalized trace on $\mathcal{P}^N\,\mathcal{H}^N$. Condition (13.8) is a law of large numbers for observables M_k^N under quantum state (13.9); it can be equivalently written as the condition

$$\lim_{N\uparrow+\infty} \mathrm{tr}^N[\mathcal{Q}_k^{N,\delta}(m_k)\,|\,\mathcal{P}^N] = 1 \tag{13.10}$$

for any $\delta > 0$ and $k \in I$. Physically, it means that all M_k^N, $k \in I$ are asymptotically dispersionless under state (13.9).

Having in mind the classical situation where the entropy counts the number of *all* microstates x such that $M^N(x) \overset{\delta}{=} m$, we are mostly interested in those concentrating sequences that are maximal in the sense of dimension counting. Hence, we define the (infinite-system, Boltzmann) entropy $s(m)$ for any $m = (m_k)_{k\in I}$ by the limit

$$s(m) = \limsup_{\mathcal{P}^N\to m} \frac{1}{N} \log \mathrm{Tr}^N[\mathcal{P}^N] \tag{13.11}$$

i.e., $s(m)$ is the largest limit point over all projections concentrating at m. Any projections $\mathcal{P}^N$ attaining the entropy $s(m)$ in the large N limit,

$$\limsup_{N\uparrow+\infty} \frac{1}{N} \log \mathrm{Tr}^N[\mathcal{P}^N] = s(m) \tag{13.12}$$

are then called *typical* projections concentrating at m.[21] Clearly, they provide a variant of the microcanonical ensemble for noncommuting observables. An example comes in Section 14.

To check that the above definition of entropy is meaningful, we first revisit the relation between the macroscopic autonomy and the H-theorem of Section 11.2 in the present quantum set-up. Second, we link our construction to the canonical construction based on maximizing the von Neumann entropy, and we prove that they are equivalent under suitable conditions.

13.2 H-theorem

As a microscopic dynamics we consider a family of automorphisms[22] $(\tau_t^N)_{t\geq 0}$ acting on the observables from $\mathcal{B}(\mathcal{H}^N)$ and having the semigroup property

$$\tau_t^N \tau_s^N = \tau_{t+s}^N\,, \qquad t,s \geq 0 \tag{13.13}$$

Denote

$$\Omega_0 = \{m \in \mathbb{R}^I;\, s(m) \geq 0\} \tag{13.14}$$

the set of all admissible macrostates. The conditions on the emergent macroscopic dynamics now have the following form, cf. Section 11.2. There are maps $(\phi_t)_{t\geq 0}$ on Ω_0 satisfying

1. *semigroup condition*: $\phi_t \circ \phi_s = \phi_{t+s}$, $t,s \geq 0$;
2. *autonomy condition*: for every $m \in \Omega_0$ there exist some typical projections $\mathcal{P}^N \to m$ concentrating at m such that for all $F \in C(\mathbb{R})$, $k \in I$, and $t \geq 0$,

$$\lim_{N\uparrow+\infty} \mathrm{tr}^N[\tau_t^N F(M_k^N) \,|\, \mathcal{P}^N] = F((\phi_t m)_k) \tag{13.15}$$

or, equivalently,

$$\lim_{\delta\downarrow 0} \lim_{N\uparrow+\infty} \mathrm{tr}^N[\tau_t^N \mathcal{Q}_k^{N,\delta}((\phi_t m)_k) \,|\, \mathcal{P}^N] = 1 \tag{13.16}$$

[21] Note a slight difference in the terminology with respect to [12] where the typical concentrating projections were rather called a microcanonical macrostate. The present terminology is closer to the one of quantum information theory.

[22] This means that $\tau_t^N(XY) = \tau_t^N(X)\,\tau^N(Y)$ for any $X, Y \in \mathcal{B}(\mathcal{H}^N)$, which is a noncommutative generalization of classical deterministic map. Physically, $(\mathcal{H}^N, \tau_t^N, \mathrm{Tr}^N)$ models a *closed* quantum dynamical system; note that $\mathrm{Tr}^N(\tau_t^N(\cdot)) = \mathrm{Tr}^N(\cdot)$ and hence Tr^N corresponds to the invariant (counting, unnormalized) classical measure.

Under these condition it is easy to prove that

$$s(\phi_t m) \geq s(\phi_s m)\,, \qquad m \in \Omega_0,\, t \geq s \geq 0 \tag{13.17}$$

Indeed, let $\mathcal{P}^N \to m$ be typical projections concentrating at m and verifying (13.15) or (13.16). Using that τ_t^N is invertible and $(\tau_t^N)^{-1}$ is again an automorphism such that $\mathrm{Tr}^N((\tau_t^N)^{-1}\cdot) = \mathrm{Tr}^N(\cdot)$, one has

$$\begin{aligned} \mathrm{tr}^N[\tau_t^N F(M_k^N)\,|\,\mathcal{P}^N] &= \frac{\mathrm{tr}^N[F(M_k^N)\,(\tau_t^N)^{-1}\mathcal{P}^N]}{\mathrm{tr}^N[(\tau_t^N)^{-1}\mathcal{P}^N]} \\ &= \mathrm{tr}^N[F(M_k^N)\,|\,(\tau_t^N)^{-1}\mathcal{P}^N] \end{aligned} \tag{13.18}$$

Hence, autonomy (13.15) implies that $(\tau_t^N)^{-1}\mathcal{P}^N$ concentrate at $\phi_t m$, $(\tau_t^N)^{-1}\mathcal{P}^N \to \phi_t m$.[23] As a result,

$$s(m) = \limsup_{N\uparrow+\infty} \frac{1}{N} \log \mathrm{Tr}^N[(\tau_t^N)^{-1}\mathcal{P}^N] \leq s(\phi_t m) \tag{13.19}$$

By combining with the semigroup property,

$$s(\phi_t m) = s(\phi_{t-s} \circ \phi_s m) \geq s(\phi_s m) \tag{13.20}$$

as claimed.

Notice that in (13.15)–(13.16) we have required the autonomy condition in the sense of a law of large numbers; compare with a much weaker assumption (11.4). A possible way how to prove the H-theorem under a weaker autonomy condition here too, might be via suitably weakening the notion of concentration and by modifying the definition of entropy; we do not discuss this issue.

13.3 Canonical Formalism

Von Neumann has introduced the entropy functional on states $\omega^N(\cdot) = \mathrm{Tr}^N[\rho^N\cdot\,]$ over $\mathcal{B}(\mathcal{H}^N)$ by

$$\mathfrak{S}(\omega^N) = -\,\mathrm{Tr}^N[\rho^N \log \rho^N] \tag{13.21}$$

For trace states on subspaces of $\mathcal{H}^N$, given as

$$\omega^N_{\mathcal{P}^N}(\cdot) = \mathrm{Tr}^N[\mathcal{P}^N\cdot\,]/\,\mathrm{Tr}^N[\mathcal{P}^N], \tag{13.22}$$

the von Neumann entropy boils down to

$$\mathfrak{S}(\omega^N_{\mathcal{P}^N}) = \log \mathrm{Tr}^N[\mathcal{P}^N] \tag{13.23}$$

[23] Note they do not have to be *typical* concentrating projections at $\phi_t m$!

In this light, entropy (13.11) can also be written as

$$s(m) = \limsup_{\mathcal{P}^N \to m} \frac{\mathfrak{S}(\omega^N_{\mathcal{P}^N})}{N} \tag{13.24}$$

A general and very successful approach in statistical physics lies in the idea that a variational principle like (13.24) can often be extended to a larger 'test space', so that (1) a new variational problem becomes easier to solve, and (2) the resulting entropy $s(m)$ can be proven to remain unchanged. This is a standard approach at least when describing thermal equilibrium, but it is often used in a similar way to describe nonequilibrium macroscopic states (sometimes then referred to as *constrained* equilibria).

To obtain the *canonical* description for a given macrostate $m \in \mathbb{R}^I$, we write $\omega^N \stackrel{1}{\to} m$ for any sequence of states satisfying

$$\lim_{N\uparrow+\infty} \omega^N(M^N_k) = m_k, \quad k \in I \text{ (convergence in mean).} \tag{13.25}$$

Analogous to (13.24), we define the *canonical* entropy,

$$s_{\rm can}(m) = \limsup_{\omega^N \stackrel{1}{\to} m} \frac{\mathfrak{S}(\omega^N)}{N} \tag{13.26}$$

Any sequence of states $(\omega^N)_{N\uparrow+\infty}$ such that $\lim_{N\uparrow+\infty} \mathfrak{S}(\omega^N)/N = s_{\rm can}(m)$ we then call *canonical states* at m.

An advantage of this formulation is that one can often find canonical states explicitly in a Gibbsian form: consider states $\omega^N_\lambda(\cdot) = {\rm Tr}^N[\rho^N_\lambda \cdot\,]$ defined as

$$\rho^N_\lambda = \frac{1}{\mathcal{Z}^N_\lambda}\, e^{N\sum_k \lambda_k M^N_k}, \qquad \mathcal{Z}^N_\lambda = {\rm Tr}^N[e^{N\sum_k \lambda_k M^N_k}] \tag{13.27}$$

with some $\lambda = (\lambda_k)_{k\in I}$. If $\lim_{N\uparrow+\infty} \omega^N_\lambda(M^N_k) = m_k$, $k \in I$, then $(\omega^N_\lambda)_{N\uparrow+\infty}$ are canonical states at m.

This easily follows from the positivity of relative entropy, see e.g. [5]: for any $\omega^N \stackrel{1}{\to} m$, $\omega^N(\cdot) = {\rm Tr}^N[\rho^N \cdot\,]$,

$$\begin{aligned} \limsup_{N\uparrow+\infty} -\frac{1}{N}\,\omega^N[\log \rho^N] &\le \limsup_{N\uparrow+\infty} -\frac{1}{N}\,\omega^N[\log \rho^N_\lambda] \\ &= \limsup_{N\uparrow+\infty} \frac{1}{N} \log \mathcal{Z}^N_\lambda - \sum_k \lambda_k m_k \\ &= \limsup_{N\uparrow+\infty} -\frac{1}{N}\,\omega^N_\lambda[\log \rho^N_\lambda] \end{aligned} \tag{13.28}$$

as claimed. It also yields the canonical entropy in the form

$$s_{\rm can}(m) = p(\lambda) - \sum_k \lambda_k m_k \tag{13.29}$$

where we have defined the 'pressure'

$$p(\lambda) = \limsup_{N\uparrow+\infty} \frac{1}{N} \log \mathcal{Z}_\lambda^N \tag{13.30}$$

13.4 Macroscopic Equivalence

By construction, $s_{\rm can}(m) \geq s(m)$. A natural question arises under what conditions both entropies are actually equal. This is a familiar problem of the equivalence of ensembles (microcanonical versus canonical in this case) on the level of entropies, however, the usual arguments, e.g. [37, 55, 19], are mostly restricted to the case of equilibrium and commuting observables (with the energy and/or the particle number as the only variables). The generalized microcanonical ensemble in the sense of Section 13.1 requires some refinement of those arguments. Below we provide some sufficient conditions for the equivalence.

Let $(\omega_\lambda^N)_{N\uparrow+\infty}$ be canonical states (13.27) with $\omega_\lambda^N \stackrel{1}{\to} m$. Assume that

1. the limit
$$p(\lambda) = \lim_{N\uparrow+\infty} \frac{1}{N} \log {\rm Tr}^N [e^{N\sum_k \lambda_k M_k^N}] \tag{13.31}$$
exists and has the derivative $\frac{{\rm d}p(\kappa\lambda)}{{\rm d}\kappa}\big|_{\kappa=1} = \sum_k \lambda_k\, m_k$;
2. for any $j \in I$, the generating function
$$q_j(\kappa) = \lim_{N\uparrow+\infty} \frac{1}{N} \log {\rm Tr}^N [e^{N\sum_k \lambda_k M_k^N} e^{\kappa N M_j^N}] \tag{13.32}$$
exists and has the derivative $\frac{{\rm d}q_j(\kappa)}{{\rm d}\kappa}\big|_{\kappa=0} = m_j$.

Under these hypotheses we will prove that

$$s(m) = s_{\rm can}(m) = \sum_{k\in I} \lambda_k m_k - p(\lambda) \tag{13.33}$$

Remark that by the Golden-Thompson inequality,[24]

$$q_j(\kappa) \geq p(\lambda + (0, \ldots, (\kappa)_j, \ldots, 0)) \tag{13.34}$$

[24] $e^{A+B} \leq e^A e^B$ for all hermitian matrices A, B.

Unless M_N^k all mutually commute, this inequality generically becomes strict and those $(q_j)_{j\in I}$ are fundamentally different from the pressure; they appear naturally when studying quantum large fluctuations, see Section 13.5.

The proof of equivalence (13.33) comes in a sequence of steps: first we show that the canonical states ω_λ^N are exponentially concentrating at m, then we construct typical projections for these states, and finally we prove that those typical projections concentrate at m too.

Exponential Concentration

By assumption, q_j exists in some interval $[-\kappa_0, \kappa_0]$, $\kappa_0 > 0$. From the spectral theorem (13.3),

$$\begin{aligned} q_j(\kappa) &= p(\lambda) + \lim_{N\uparrow+\infty} \frac{1}{N} \log \omega_\lambda^N [e^{\kappa N M_j^N}] \\ &= p(\lambda) + \lim_{N\uparrow+\infty} \frac{1}{N} \log \int_{\mathbb{R}} \omega_\lambda^N [\mathcal{Q}_j^N(\mathrm{d}z)]\, e^{\kappa z N} \\ &\equiv p(\lambda) + \lim_{N\uparrow+\infty} \frac{1}{N} \log \int_{\mathbb{R}} \nu_j^N(\mathrm{d}z)\, e^{\kappa z N} \end{aligned} \tag{13.35}$$

where we have introduced the (classical) probability measures ν_j^N as the distribution of M_j^N under states ω_λ^N. Denote the last term as $\psi_j(\kappa)$, and fix some $\delta > 0$. One has the estimate

$$\int_{\mathbb{R}} \nu_j^N(\mathrm{d}z)\, e^{\kappa z N} \geq e^{\kappa(m_j+\delta)N} \nu_j^N[z \geq m_j + \delta] \tag{13.36}$$

which implies, by the existence of the limiting generating function,

$$\limsup_{N\uparrow+\infty} \frac{1}{N} \log \nu_j^N[z \geq m_j + \delta] \leq \psi_j(\kappa) - \kappa(m_j + \delta) \tag{13.37}$$

for all $0 \leq \kappa \leq \kappa_0$. Since $\frac{\mathrm{d}\psi_j}{\mathrm{d}\kappa}|_{\kappa=0} = m_j$, there exists $\kappa_1 = \kappa_1(\delta)$, $0 < \kappa_1 \leq \kappa_0$ such that $\psi_j(\kappa_1) \leq \kappa_1 m_j + \frac{\kappa_1 \delta}{2}$. Hence,

$$\limsup_{N\uparrow+\infty} \frac{1}{N} \log \nu_j^N[z \geq m_j + \delta] \leq -\frac{\kappa_1 \delta}{2} \tag{13.38}$$

Combining with an analogous argument for $\nu_j^N[z \leq m_j - \delta]$, we arrive at the bound

$$\omega_\lambda^N[Q_j^{N,\delta}(m_j)] \geq 1 - e^{-C_j(\delta)N} \tag{13.39}$$

valid for all $\delta > 0$ and $N \geq N_j(\delta)$, with some $C_j(\delta) > 0$ and $N_j(\delta)$; the $\mathcal{Q}_j^{N,\delta}(m_j)$ is given by (13.4).

This in particular implies that states ω_λ^N are concentrating at $m(\lambda)$ in the sense of a law of large numbers analogous to (13.10); moreover, the concentration is exponentially fast. Note that the above argument is similar to the construction of large deviation upper bounds, cf. any textbook on the large deviation theory, e.g. [9, 23].

In an analogous way we exploit assumption (1) on the pressure. This time we consider the observable $\sum_k \lambda_k M_k^N$ and denote by $\bar{\mathcal{Q}}^N$ the corresponding projection-valued measure, i.e., such that

$$F\big(\sum_k \lambda_k M_k^N\big) = \int_{\mathbb{R}} \bar{\mathcal{Q}}^N(\mathrm{d}z)\, F(z)\,, \qquad F \in C(\mathbb{R}) \tag{13.40}$$

Repeating the arguments (13.35)–(13.39), we get the result

$$\omega_\lambda^N[\bar{Q}^{N,\delta}] \geq 1 - e^{-\bar{C}(\delta)N} \tag{13.41}$$

with

$$\bar{Q}^{N,\delta} = \int_{\mathbb{R}} \bar{Q}^N(\mathrm{d}z)\, \chi\big(\sum_k \lambda_k m_k - \delta \leq z \leq \sum_k \lambda_k m_k + \delta\big) \tag{13.42}$$

valid again for all $\delta > 0$, $N \geq \bar{N}(\delta)$, with some $\bar{C}(\delta) > 0$ and $\bar{N}(\delta)$.

Typical Projections

From (13.41) there is a sequence $\delta_N \downarrow 0$ such that the projections $\mathcal{P}^N = \bar{\mathcal{Q}}^{N,\delta_N}$ satisfy

$$\lim_{N\uparrow+\infty} \omega_\lambda^N[\mathcal{P}^N] = 1 \tag{13.43}$$

By construction one has the operator inequalities

$$\mathcal{P}^N\big(\sum_k \lambda_k m_k - \delta_N\big) \leq \mathcal{P}^N \sum_k \lambda_k M_k^N \leq \mathcal{P}^N\big(\sum_k \lambda_k m_k + \delta_N\big) \tag{13.44}$$

which yield the upper bound

$$\mathrm{Tr}^N[\mathcal{P}^N] = \omega_\lambda^N[(\rho_\lambda^N)^{-1}\mathcal{P}^N] \leq \mathcal{Z}_\lambda^N e^{-N\left(\sum_k \lambda_k m_k - \delta_N\right)} \omega_\lambda^N[\mathcal{P}^N] \tag{13.45}$$

and the lower bound

$$\mathrm{Tr}^N[\mathcal{P}^N] \geq \mathcal{Z}_\lambda^N e^{-N\left(\sum_k \lambda_k m_k + \delta_N\right)} \omega_\lambda^N[\mathcal{P}^N] \tag{13.46}$$

Using (13.43) and (13.29), this proves[25]

$$\lim_{N\uparrow+\infty} \frac{1}{N} \log \mathrm{Tr}^N[\mathcal{P}^N] = p(\lambda) - \sum_k \lambda_k m_k = s_{\mathrm{can}}(m) \tag{13.47}$$

As soon as we prove that projections $\mathcal{P}^N$ are concentrating at m (see the next section), the last equation simply means that $s(m) \geq s_{\mathrm{can}}(m)$. Since the opposite inequality is obvious, we arrive at (13.33) as claimed.

The arguments used in this section are well known in both classical and quantum information theory, and projections $\mathcal{P}^N$ satisfying (13.43) and (13.47) are usually called *typical (sequence of) projections*. Their existence under mild assumptions for a large class of models is a subject of the Shannon-McMillan(-Breiman) theorem, see e.g. [3] and references therein. For a nice overview of the principles of quantum information theory see [8].

Concentration of Typical Projections

To finish the proof we need to show that $\mathcal{P}^N$ as constructed in the last section concentrate at m. The following is true for any $Y^N \geq 0$:

$$\begin{aligned}
\omega_\lambda^N[Y^N] &= \mathrm{Tr}^N[(\rho_\lambda^N)^{\frac{1}{2}} Y^N (\rho_\lambda^N)^{\frac{1}{2}}] \\
&\geq \mathrm{Tr}^N[\mathcal{P}^N (\rho_\lambda^N)^{\frac{1}{2}} Y^N (\rho_\lambda^N)^{\frac{1}{2}} \mathcal{P}^N] \\
&= \mathrm{Tr}^N[(Y^N)^{\frac{1}{2}} \mathcal{P}^N \rho_\lambda^N (Y^N)^{\frac{1}{2}}] \\
&\geq \frac{1}{\mathcal{Z}_\lambda^N} e^{N\left(\sum_k \lambda_k m_k - \delta_N\right)} \mathrm{Tr}^N[\mathcal{P}^N] \ \mathrm{tr}^N[Y^N \,|\, \mathcal{P}^N] \\
&\geq e^{-2N\delta_N} \omega^N[\mathcal{P}^N] \ \mathrm{tr}^N[Y^N \,|\, \mathcal{P}^N]
\end{aligned} \tag{13.48}$$

where we have used inequalities (13.44) and (13.46). Take now $Y^N = 1 - \mathcal{Q}_j^{N,\epsilon}(m_j)$ and use the exponential concentration property of ω_λ^N, inequality (13.39); one obtains

$$1 - \mathrm{tr}^N[\mathcal{Q}_j^{N,\epsilon}(m_j) \,|\, \mathcal{P}^N] \leq e^{-(C_j(\epsilon) - 2\delta_N)N} (\omega_\lambda^N[\mathcal{P}^N])^{-1} \tag{13.49}$$

for any $\epsilon > 0$ and $N \geq N_j(\epsilon)$, which immediately gives[26]

[25] Note that one only needs that $\lim_{N\uparrow+\infty} \frac{1}{N} \log \mathrm{Tr}^N[\mathcal{P}^N] = 0$. In particular, the assumption on the differentiability of the pressure is convenient but far from necessary; see also a comment below.

[26] Note it actually yields an exponential concentration, even under that weaker assumption $\lim_{N\uparrow+\infty} \frac{1}{N} \log \mathrm{Tr}^N[\mathcal{P}^N] = 0$.

$$\lim_{N\uparrow+\infty} \mathrm{tr}^N[\mathcal{Q}_j^{N,\epsilon}(m_j)\,|\,\mathcal{P}^N] = 1 \tag{13.50}$$

Repeating for all $j \in I$, this proves $\mathcal{P}^N \to m$.

13.5 Towards Quantum Large Deviations

Using the notation of Section 13.4, the Gärtner-Ellis theorem of (classical) large deviations, [9, 23], teaches us that whenever the generating function $q_j \in C^1(\mathbb{R})$ is differentiable and strictly convex, one has the law

$$\lim_{\delta\downarrow 0} \lim_{N\uparrow\infty} \frac{1}{N} \log \nu_j^N[z \overset{\delta}{=} \tilde{m}_j] = -I_j(\tilde{m}_j) \tag{13.51}$$

for any $\tilde{m}_j$ such that $\tilde{m}_j = \frac{\mathrm{d}q_j}{\mathrm{d}\kappa}\big|_{\kappa=\kappa(\tilde{m}_j)}$ for some (unique by assumption) $\kappa(\tilde{m}_j)$, and with the rate function I_j being the Legendre transform

$$\begin{aligned} I_j(\tilde{m}_j) &= \sup_\kappa \big\{ \kappa\, \tilde{m}_j - \lim_{N\uparrow+\infty} \frac{1}{N} \log \omega_\lambda^N \big[e^{\kappa N M_j^N} \big] \big\} \\ &= \sup_\kappa [\kappa\, \tilde{m}_j - q_j(\kappa) + p(\lambda)] \\ &= \kappa(\tilde{m}_j)\, \tilde{m}_j - q_j(\kappa(\tilde{m}_j)) + p(\lambda) \end{aligned} \tag{13.52}$$

(Naturally, for m such that $\omega_\lambda^N \overset{1}{\to} m$ one has $\kappa(m_j) = 0$ and $I_j(m_j) = 0$.) In terms of the canonical states ω_λ^N, (13.51) becomes simply

$$\lim_{\delta\downarrow 0} \lim_{N\uparrow+\infty} \frac{1}{N} \log \omega_\lambda^N[\mathcal{Q}_j^{N,\delta}(\tilde{m}_j)] = -I_j(\tilde{m}_j) \tag{13.53}$$

This is an exponential law for the outcomes of measurements of the observables M_j^N, upon the canonical states ω_λ^N. This gives an interpretation to q_j as the corresponding generating function.

The existence and differentiability of q_j gets nontrivial whenever the observables M_k^N are more complicated than just something like the spatial averages of one-site observables over a lattice (a simple example are the observables M_1^N, M_2^N, M_3^N in the quantum Kac model, Section 14). In the usual context of quantum lattice models, no general argument is known even for the existence of $q_j(\kappa)$, which is in contrast to the case of pressure $p(\lambda)$ where the situation is rather well understood, [55, 24, 5]. For some partial results about the existence of q_j in the so called high-temperature regime see [49, 53]; the differentiability is studied in [49].

By means of the Varadhan lemma, [9], formula (13.53) can be equivalently written as

$$\lim_{N\uparrow+\infty} \frac{1}{N} \log \omega_\lambda^N [e^{NF(M_j^N)}] = \sup_z \{F(z) - I_j(z)\} \tag{13.54}$$

for any $F \in C(\mathbb{R})$, for simplicity assumed to be bounded from above. This form provokes still another related question, namely the asymptotic limit

$$\lim_{N\uparrow+\infty} \frac{1}{N} \log \mathrm{Tr}^N [e^{N\left(\sum_k \lambda_k M_k^N + F(M_j^N)\right)}] \tag{13.55}$$

Although this is likely not directly related to the quantum fluctuations, such formulas appear naturally when studying lattice models with a combination of short-range and long-range interactions. Some authors consider *this* formulation as a genuine problem of quantum large deviations; see e.g. [51] where the authors show the above limit to exist in the case of M_k^N being averages over one-site spin observables.[27] They prove the following variational principle:

$$(13.55) = \sup_z \{F(z) - I'_j(z)\}\,, \qquad I'_j(z) = \sup_\kappa \{\kappa\, z - q'_j(\kappa)\} \tag{13.56}$$

$$q'_j(\kappa) = \lim_{N\uparrow+\infty} \frac{1}{N} \log \mathrm{Tr}^N \left[e^{N\left(\sum_k \lambda_k M_k^N + \kappa M_j^N\right)}\right] \tag{13.57}$$

which is similar to (13.51)–(13.52) up to the modified generating function q'.

A general and systematic quantum large deviation theory is lacking, however, and remains an interesting open question. Possibly even more ambitious, both physically and mathematically, would be the problem of correlated large fluctuations for noncommuting macroscopic observables. Some ideas on this issue can be found in [2]; also the present construction of generalized microcanonical ensembles, Section 13.1, seems related to this problem.

14 Example: Quantum Kac Ring

This is a quantum extension of the Kac ring model of Section 10, introduced and studied in [15, 11]. Consider a ring $\Lambda = \{1, \ldots, N\}$ again, and associate with each site i a quantum spin $\eta(i) \in \mathbb{C}^2$ and a classical variable $g(i) \in \{1, 0\}$ that indicates the presence respectively the absence of a scatterer. The state space of the model is hence $\mathcal{H}^N \times K^N$ with Hilbert space $\mathcal{H}^N = \mathbb{C}^{2N}$ (spins) and classical space $K^N = \{0, 1\}^N$ (scatterers).

[27] This in particular means that the canonical states ω_λ^N are product states.

14.1 Macroscopic Description

As macroscopic observables we consider the operators

$$M_\alpha^N = \frac{1}{N}\sum_{i=1}^{N}\sigma_\alpha(i)\,, \qquad \alpha = 1,2,3 \tag{14.1}$$

where $\sigma_\alpha(i)$ are copies at site i of the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{14.2}$$

representing three components of the local 'magnetization', and

$$M_0^N(g) = \frac{1}{N}\sum_{i=1}^{N} g(i) \tag{14.3}$$

the density of scatterers. By construction, $[M_1^N, M_2^N] = \frac{1}{N}M_3^N$ (and cyclic permutations). The classical (= commutative) case is restored by keeping e.g. M_0^N and M_3^N as the only macroobservables.

It is sometimes convenient to embed K^N in $\mathbb{C}^{2N}$ and to utilize a compact notation for both operators on $\mathcal{H}^N$ and classical functions on K^N. In this sense we speak below about states over $\mathcal{H}^N \times K^N$, and we use the shorthand $\hat{\mathrm{Tr}}^N = \sum_{g\in K^N} \mathrm{Tr}^N$.

In the canonical framework,

$$\begin{aligned} \omega_\lambda^N(\cdot) &= \frac{1}{\mathcal{Z}_\lambda^N}\,\hat{\mathrm{Tr}}^N[e^{N\sum_{\alpha=0}^{3}\lambda_\alpha M_\alpha^N}\cdot] \\ &= e^{-Np(\lambda)}\hat{\mathrm{Tr}}^N\big[e^{\sum_{i=1}^{N}\left(\lambda_0 g(i)+\sum_{\alpha=1}^{3}\lambda_\alpha\sigma_\alpha(i)\right)}\cdot\big] \end{aligned} \tag{14.4}$$

are product canonical states, and the pressure is

$$p(\lambda) = \frac{1}{N}\log\hat{\mathrm{Tr}}^N[e^{N\sum_{\alpha=0}^{3}\lambda_\alpha M_\alpha^N}] = \log 2[(1+e^{\lambda_0})\cosh|\boldsymbol{\lambda}|] \tag{14.5}$$

with the shorthands $\boldsymbol{\lambda} = (\lambda_1,\lambda_2,\lambda_3)$ and $|\boldsymbol{\lambda}| = (\lambda_1^2+\lambda_2^2+\lambda_3^2)^{\frac{1}{2}}$. Further, $\omega_\lambda^N \overset{1}{\to} m$ where

$$m_0 = \frac{\partial p}{\partial\lambda_0} = (1+e^{-\lambda_0})^{-1}\,, \quad m_\alpha = \frac{\partial p}{\partial\lambda_\alpha} = \frac{\lambda_\alpha}{|\boldsymbol{\lambda}|}\tanh|\boldsymbol{\lambda}|\,,\ \alpha = 1,2,3 \tag{14.6}$$

The canonical entropy is then, $m = (m_0, \mathbf{m})$,

$$s_{\rm can}(m) = p(\lambda) - \sum_{\alpha=0}^{3} \lambda_\alpha m_\alpha$$
$$= \begin{cases} -\frac{1+|\mathbf{m}|}{2}\log\frac{1+|\mathbf{m}|}{2} - \frac{1-|\mathbf{m}|}{2}\log\frac{1-|\mathbf{m}|}{2} - m_0\log m_0 \\ \qquad -(1-m_0)\log(1-m_0) \qquad \text{if } |\mathbf{m}|<1,\ 0<m_0<1 \\ -\infty \qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise,} \end{cases} \tag{14.7}$$

cf. the classical case, (10.11).

To obtain a microcanonical description in the sense of Section 13.1, we associate with any macroscopic state $m = (m_0, \mathbf{m})$ the *modified* macroscopic observable $(M_0^N, \bar{M}^N)$,

$$\bar{M}^N = \sum_{\alpha=1}^{3} \frac{m_\alpha}{|\mathbf{m}|} M_\alpha^N = \frac{1}{N}\sum_{i=1}^{N} \bar{\sigma}(i)\,, \qquad \bar{\sigma} = \frac{\mathbf{m}}{|\mathbf{m}|}\cdot\boldsymbol{\sigma} \tag{14.8}$$

and the modified macrostate $(m_0, |\mathbf{m}|)$.

Since $\bar{\sigma}$ is unitarily equivalent to e.g. σ_3, i.e., $\bar{\sigma} = W\sigma_3 W^\dagger$ with some $W^\dagger = W^{-1}$, we are back at the classical (commutative) situation. Denoting by $\bar{\mathcal{Q}}^N(\mathrm{d}z)$ the projection-valued measure for $(M_0^N, \bar{M}^N)$, one easily checks that any[28] $\bar{\mathcal{Q}}^{N,\delta_N}(m_0, |\mathbf{m}|)$ such that $\delta_N \downarrow 0$, are concentrating projections at $(m_0, |\mathbf{m}|)$. Moreover, if $N^{\frac{1}{2}}\delta_N \uparrow +\infty$ then these are *typical* concentrating projections at $(m_0, |\mathbf{m}|)$ and the entropy is, as essentially can be read off from the classical formula (10.11),

$$s(m_0, |\mathbf{m}|) = \lim_{N\uparrow+\infty} \frac{1}{N}\log \hat{\rm Tr}^N[\bar{\mathcal{Q}}^{N,\delta_N}(m_0, |\mathbf{m}|)] = s_{\rm can}(m_0, \mathbf{m}) \tag{14.9}$$

In the last step, we need to show that $\bar{\mathcal{Q}}^{N,\delta_N}(m_0, |\mathbf{m}|)$ are also concentrating at $m = (m_0, \mathbf{m})$, that is the macrostate under the *original* (noncommuting family of) macroscopic observables M^N. This can be proven by essentially repeating the argument of Section 13.4; we leave it to reader as an exercise. As a result, those $\bar{\mathcal{Q}}^{N,\delta_N}(m_0, |\mathbf{m}|)$ are typical projections concentrating at $m = (m_0, \mathbf{m})$.

14.2 Microscopic Dynamics

To model the scattering of quantum spins (represented by vector η) on the binary variable g, consider a unitary matrix V on $\mathbb{C}^2$,

$$V = e^{i\mathbf{h}\cdot\boldsymbol{\sigma}}\,, \qquad \mathbf{h} = (h_1, h_2, h_3) \tag{14.10}$$

[28] The notation is the same as in Section 13.1.

Let the dynamics on $\mathcal{H}^N \times K^N$ be given as, cf. (10.1),

$$U^N(\eta; g) = \big(g(N)\, V\eta(N) + (1 - g(N))\, \eta(N), \\ g(1)\, V\eta(1) + (1 - g(1))\, \eta(1), \dots, \\ \dots, g(N-1)\, V\eta(N-1) + (1 - g(N-1))\, \eta(N-1);\, g\big) \quad (14.11)$$

extended to a unitary operator in the quantum sector by linearity. The associated automorphisms are then

$$\tau_t^N(\cdot) = (U^N)^{-t} \cdot (U^N)^t \quad (14.12)$$

14.3 Macroscopic Dynamics

Let us start with a heuristic argument in the spirit of Boltzmann's Stosszahlansatz. Any macrostate $m = (m_0, \mathbf{m})$ can be associated with the quantum state of a *single* 'effective' quantum spin, via the 2×2 density matrix

$$\nu = \frac{1}{2}(\mathbb{1} + \mathbf{m} \cdot \boldsymbol{\sigma})\,, \qquad \mathrm{Tr}[\nu\, \sigma_\alpha] = m_\alpha\,,\ \alpha = 1, 2, 3 \quad (14.13)$$

Each time step the effective spin either meets a scatterer (with probability m_0) or not (with probability $1 - m_0$). Hence, its evolution is presumably $\nu \mapsto \nu_t = \bar{\phi}^t(\nu)$,

$$\bar{\phi}(\nu) = m_0 V \nu V^\dagger + (1 - m_0)\, \nu \quad (14.14)$$

by construction enjoying the semigroup property. Using (14.10) and (14.13), this can be explicitly written as the evolution on macrostates: $m_{t+1} = \phi(m_t)$ where

$$\phi(m_0, \mathbf{m}) = (m_0,\ \mathbf{m} - 2m_0[(\mathbf{n} \times \mathbf{m}) \sin|\mathbf{h}| \cos|\mathbf{h}| - \mathbf{n} \times (\mathbf{n} \times \mathbf{m}) \sin^2|\mathbf{h}|]) \quad (14.15)$$

with the notation $\mathbf{n} = \mathbf{h}/|\mathbf{h}|$. One easily checks that $\mathbf{h} \cdot \mathbf{m}$ is invariant under ϕ; the evolution can be visualized as a spiral motion in the plane perpendicular to $\mathbf{n}$. Provided that $|\mathbf{h}| \neq 0, \pi, 2\pi, \dots$ and $m_0 \in (0, 1)$,

$$\lim_{t \uparrow +\infty} \phi^t(m_0, \mathbf{m}) = (m_0, (\mathbf{n} \cdot \mathbf{m})\, \mathbf{n}) \quad (14.16)$$

and the relaxation is exponentially fast. The monotonicity of the entropy $s(m_t)$ can also be easily verified.

A rigorous argument showing that the above heuristics is indeed correct was given in [10], employing a strategy similar to that for the classical Kac model, Section 10.2. The result reads that for a large class of (sequences of) states $\omega^N \stackrel{1}{\to} m$, including in particular

- the (microcanonical) states $\mathrm{tr}^N[\,\cdot\,|\,\bar{\mathcal{Q}}^{N,\delta_N}(m,|\mathbf{m}|)]$ under those typical concentrating projections at m constructed in Section 14.1;
- the canonical states ω^N_λ from (14.4);

one has the law of large numbers:

$$\omega^N[\tau^N_t F(M^N_\alpha)] = F((\phi_t m)_\alpha)\,, \qquad \alpha = 1,2,3 \tag{14.17}$$

for all $F \in C(\mathbb{R})$ and with ϕ given by (14.15). Hence, one verifies the autonomy condition (13.15).

14.4 Exercise

Consider (M^N_0, M^N_3) as a new macroscopic observable. Check that the data (m_0, m_3) are macroscopically equivalent with $(m_0, 0, 0, m_3)$ for the original 'full' macroscopic observable M^N; therefore the autonomy just follows from (14.17). Calculate the entropy $s(m_0, m_3)$ and show that it oscillates as a function of time. How can this apparent failure of the H-theorem be explained?

15 Concluding Remarks

The text has discussed some newer and some older issues of nonequilibrium physics. Main emphasis has been on fluctuations and on the relation between entropy and irreversibility. One could say that everything has been an exploration of the idea that entropy production is a measure of irreversibility. Some central identities have been (4.7), (7.5), (10.25) and (11.9) which all point to the deep connection between source terms of time-reversal breaking and statistical thermodynamic quantities. They go beyond standard irreversible thermodynamics because fluctuations play an essential role here. As known since long, the deviations of thermodynamic behavior are important in the very understanding of its microscopic origin. These relations go also beyond the standard schemes as they are not perturbative and they do not require linear approximations or closeness to equilibrium.

Nevertheless there is also a sense in which all that has been attempted here does remain very close to the standard perspective. We

do not mean only that there is not really much fundamentally new since Boltzmann's statistical interpretation of entropy. It is true that progress has been very slow and we have been writing mostly from the point of view of the rear-guard, dissecting arguments and explanations that have been won long before. What we do have in mind however is that the theory so far remains very much restricted to direct comparisons with equilibrium. The obsession with time has mostly been an interest in the passing away of structure, of deleting memory and of ending in equilibrium—all the time centering around the second law of thermodynamics, and often applying Markovian schemes or justifying molecular chaos. We hope that the lectures that are summarized in the preceding sections have indeed clarified some of these issues, but we do not want to leave the reader without trying to provoke some feeling of totally different directions.

The most sensational instances of nonequilibrium physics are probably not to be found in the problem of relaxation to equilibrium nor in the installation of nonequilibrium via standard thermodynamic forces for which the linear response theory appears to be working well even quite far from equilibrium. What needs to be understood is the constructive role of fluctuations far away from equilibrium. For example, understanding nonequilibrium aspects in life processes be it for molecular motors or for the problem of protein folding, requires fundamental studies in reaction-rate theory. Ratchet mechanisms and the physics of transport and dissipation on very small scales must be part of it also. Nonequilibrium issues that are related to macroscopic structure (even on cosmic scales), to pattern formation and to the organization of robust steady behavior are mind-boggling, but one has to open them also via the methods and the traditions of mathematical statistical physics when one wants its role to go further than "simplification and reduction of the results of previous investigations to a form in which the mind can grasp them."[29]

Acknowledgment K. N. acknowledges the support from the project AVOZ10100520 in the Academy of Sciences of the Czech Republic. The work of B. M. S. has been supported by a short term post doctoral grant at the Institute of Theoretical Physics, K.U.Leuven; Grant of K.U.Leuven – PDM/06/116 and Grant of Georgian National Science Foundation – GNSF/ST06/4-098.

[29] J. C. Maxwell, in: *On Faraday's lines of forces.*

References

1. L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, Macroscopic fluctuation theory for stationary non-equilibrium states, *J. Stat. Phys.* **107**, 635–675 (2002).
2. I. Bjelaković, J. D. Deuschel, T. Krüger, S. Seiler, R. Siegmund-Schultze, A. Szkola, A quantum version of Sanov's theorem, *Commun. Math. Phys.* **260**, 659–671 (2005).
3. I. Bjelaković, J. D. Deuschel, T. Krüger, S. Seiler, R. Siegmund-Schultze, A. Szkola, The Shannon-McMillan theorem for ergodic quantum lattice systems, *Invent. Math.* **155** (1), 203–202 (2004).
4. A. Bovier, F. den Hollander and F. Nardi, Sharp asymptotics for Kawasaki dynamics on a finite box with open boundary conditions, *Prob. Th. Rel. Fields.* **135**, 265–310 (2006).
5. O. Bratelli, D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics, Volume 2*, Springer, Berlin (1979).
6. J. Bricmont, Science of chaos or chaos in science, In: The Flight from Science and Reason, *Ann. N.Y. Academy of Science*, **79**, 731 (1996).
7. G. E. Crooks, Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems, *J. Stat. Phys.* **90**, 1481 (1998).
8. N. Datta, Quantum entropy and quantum information, In: *Les Houches, Session LXXXIII, 2005*, Eds. A. Bovier, F. Dunlop, F. den Hollander, A. van Enter, and J. Dalibard, pp. 395–465 (2006).
9. A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*, Jones and Barlett Publishers, Boston (1993).
10. W. De Roeck, T. Jacobs, C. Maes, and K. Netočný, An extension of the Kac ring model, *J. Phys. A: Math. Gen.* **36**, 1–13 (2003).
11. W. De Roeck, C. Maes, and K. Netočný, H-theorems from macroscopic autonomous equations, *J. Stat. Phys.* **123**, 571–584 (2006).
12. W. De Roeck, C. Maes, and K. Netočný, Quantum macrostates, equivalence of ensembles and an H-theorem, *J. Math. Phys.* **47**, 073303 (2006).
13. B. Derrida, J. L. Lebowitz, and E. R. Speer, Free Energy Functional for Nonequilibrium Systems: An Exactly Soluble Case, *Phys. Rev. Lett.* **87**, 150601 (2001); *J. Stat. Phys.* **107**, 599 (2002).
14. M. D. Donsker, S. R. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, I., *Comm. Pure Appl. Math.* **28**:1–47 (1975).
15. M. Dresden, F. Feiock, Models in nonequilibrium quantum statistical mechanics, *J. Stat. Phys.* **4**, 111–173 (1972).
16. D. E. Evans, E. G. D. Cohen, and G. P. Morriss, Probability of second law violations in steady flows, *Phys. Rev. Lett.* **71**, 2401–2404 (1993). `cond-mat/9908420`

17. G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in nonequilibrium statistical mechanics, *Phys. Rev. Lett.* **74**, 2694–2697 (1995); Dynamical ensembles in stationary states, *J. Stat. Phys.* **80**, 931–970 (1995).
18. P. L. Garrido, S. Goldstein and J. L. Lebowitz, The Boltzmann entropy for dense fluids not in local equilibrium, *Phys. Rev. Lett.* **92**, 050602 (2003).
19. H.-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter, New York (1988).
20. S. Goldstein and J. L. Lebowitz, On the (Boltzmann) entropy of nonequilibrium systems, *Physica D* **193**, 53–56 (2004).
21. S. R. de Groot and P. Mazur, *Non-equilibrium Thermodynamics*, North Holland Publishing Company (1962).
22. P. Hänggi, P. Talkner, and M. Borkovec, Reaction rate theory: fifty years after Kramers, *Rev. of Mod. Phys.* **62**, 251–341 (1990).
23. F. den Hollander, *Large Deviations*, Field Institute Monographs, Providence, Rhode Island (2000).
24. R. Israel, *Convexity in the Theory of Lattice Gases*, Princeton University Press (1978).
25. C. Jarzynski, Nonequilibrium equality for free energy differences, *Phys. Rev. Lett.* **78**, 2690–2693 (1997).
26. E. T. Jaynes, The second law as physical fact and as human inference (1990) (unpublished), download from `http://www.etjaynescenter.org/bibliography.shtml`.
27. E. T. Jaynes, The evolution of Carnot's principle, In: *Maximum-Entropy and Bayesian Methods in Science and Engineering, 1*, Eds. G. J. Erickson and C. R. Smith, Kluwer, Dordrecht, pp. 267–281 (1988).
28. E. T. Jaynes, Gibbs vs Boltzmann entropies, In: *Papers on Probability, Statistics and Statistical Physics*, Reidel, Dordrecht (1983); Am. J. Phys. **33**, 391–398 (1965).
29. M. Kac. *Probability and Related Topics in Physical Sciences*, Interscience Publishers Inc., New York (1959).
30. C. Kipnis and C. Landim, *Scaling Limits of Interacting Particle Systems*, Springer-Verlag, Berlin (1999).
31. R. Kubo, K. Matsuo, and K. Kitahara, Fluctuation and Relaxation of Macrovariables, *J. Stat. Phys.* **9**, 51–95 (1973).
32. O. E. Lanford III, Time evolution of large classical systems, *Lecture Notes in Phys.* **38**, 1–111 (1975); The hard sphere gas in the Boltzmann-Grad limit, *Physica* **106A**, 70–76 (1981).
33. J. L. Lebowitz, Microscoic origins of irreversible macroscopic behavior, *Physica A* **263**, 516–527 (1999).
34. J. L. Lebowitz, C. Maes, and E. R. Speer, Statistical mechanics of probabilistic cellular automata, *J. Stat. Phys.* **59** 117–170 (1990).
35. J. L. Lebowitz and H. Spohn, A Gallavotti-Cohen type symmetry in the large deviation functional for stochastic dynamics, *J. Stat. Phys.* **95**, 333–365 (1999).
36. T. M. Liggett, *Interacting Particle Systems*, Springer (1985).

37. R. Lima, Equivalence of ensembles in quantum lattice systems: states, *Commun. Math. Phys.* **24**, 180–192 (1972).
38. C. Maes, On the origin and the use of fluctuation relations for the entropy, Poincaré Seminar 2003, Eds. J. Dalibard, B. Duplantier and V. Rivasseau, Birkhäuser (Basel), pp. 145–191 (2004).
39. C. Maes. Fluctuation theorem as a Gibbs property, *J. Stat. Phys.* **95**, 367–392 (1999).
40. C. Maes and K. Netočný, Time-reversal and entropy, *J. Stat. Phys.* **110**, 269–310 (2003).
41. C. Maes and K. Netočný, Static and dynamical nonequilibrium fluctuations, `cond-mat/0612525`.
42. C. Maes, K. Netočný, and M. Verschuere, Heat conduction networks, *J. Stat. Phys.* **111**, 1219–1244 (2003).
43. C. Maes, F. Redig, and A. Van Moffaert, On the definition of entropy production via examples, *J. Math. Phys.* **41**, 1528–1554 (2000).
44. C. Maes, F. Redig, and A. Van Moffaert, The restriction of the Ising model to a layer, *J. Stat. Phys.* **96**, 69–107 (1999).
45. C. Maes and K. Netočný, The canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states, *Europhysics Letters* **82**, 30003 (2008).
46. C. Maes, K. Netočný and B. Wynants, On and beyond entropy production; the case of Markov jump processes, *Markov Processes Relat. Fields* **14**, 445–464 (2008).
47. C. Maes and K. Netočný, Minimum entropy production principle from a dynamical fluctuation law, *J.Math.Phys.* **48**, 053306 (2007).
48. C. Maes, K. Netočný and B. Wynants, Steady state statistics of driven diffusions, *Physica* A **387**, 2675–2689 (2008).
49. K. Netočný and F. Redig, Large deviations for quantum spin systems, *J. Stat. Phys.* **117**, 521–547 (2004).
50. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Translated form the German edition by R. T. Beyer, Princeton University Press, Princeton (1955).
51. D. Petz, G. A. Raggio, and A. Verbeure, Asymptotics of Varadhan-type and the Gibbs variational principle, *Comm. Math. Phys.* **121**, 271–282 (1989).
52. I. Prigogine, *Introduction to Non-Equilibrium Thermodynamics*, Wiley-Interscience, New York (1962).
53. M. Lenci, L. Rey-Bellet, Large deviations in quantum lattice systems: one-phase region, *J. Stat. Phys.*, **119**, 715–746 (2005).
54. D. Ruelle, Smooth dynamics and new rheoretical ideas in nonequilibrium statistical mechanics, *J. Stat. Phys.* **95**, 393–468 (1999).
55. B. Simon, *The Statistical Mechanics of Lattice Gases*, Princeton Univ. Press, Princeton (1993).
56. H. Spohn, *Large Scale Dynamics of Interacting Particles*, Springer Verlag, Heidelberg (1991).

Facilitated Spin Models: Recent and New Results

N. Cancrini[1], F. Martinelli[2], C. Roberto[3*], and C. Toninelli[4*]

[1] Dip. Matematica Univ.l'Aquila, 1-67100 L'Aquila, Italy
nicoletta.cancrini@roma1.infn.it
[2] Dip. Matematica, Univ. Roma Tre, Largo S.L. Murialdo 00146, Roma, Italy martin@mat.uniroma3.it
[3] Universite Paris-est, L.A.M.A. UMR 8050, 5 bd Descartes, 77454 Marne-la-Vallée France
cyril.roberto@univ-mlv.fr
[4] Laboratoire de Probabilités et Modèles Alèatoires CNRS-UMR 7599 Universités Paris VI-VII 4, Place Jussieu F-75252 Paris Cedex 05 France
ctoninel@ccr.jussieu.fr

Summary. Facilitated or kinetically constrained spin models (KCSM) are a class of interacting particle systems reversible w.r.t. to a simple product measure. Each dynamical variable (spin) is re-sampled from its equilibrium distribution only if the surrounding configuration fulfills a simple local constraint which *does not involve* the chosen variable itself. Such simple models are quite popular in the glass community since they display some of the peculiar features of glassy dynamics, in particular they can undergo a dynamical arrest reminiscent of the liquid/glass transition. Due to the fact that the jumps rates of the Markov process can be zero, the whole analysis of the long time behavior becomes quite delicate and, until recently, KCSM have escaped a rigorous analysis with the notable exception of the East model. In these notes we will mainly review several recent mathematical results which, besides being applicable to a wide class of KCSM, have contributed to settle some debated questions arising in numerical simulations made by physicists. We will also provide some interesting new extensions. In particular we will show how to deal with interacting models reversible w.r.t. to a high temperature Gibbs measure and we will provide a detailed analysis of the so called one spin facilitated model on a general connected graph.

* This work was partially supported by gdre CNRS-INdAM 224, GREFIMEFI.

[1] The material presented here is an expanded version of a series of lectures delivered by F. Martinelli at the Prague summer school 2006 on *Mathematical Statistical Mechanics*.

R. Kotecký (ed.), *Methods of Contemporary Mathematical Statistical Physics*, Lecture Notes in Mathematics 1970, DOI: 10.1007/978-3-540-92796-9_7,

Contents

Key words: Glauber dynamics, spectral gap, kinetically constrained models, dynamical phase transition, glass transition.

1 Introduction and Motivations

Consider the following simple interacting particle system. At each site of the lattice $\mathbb{Z}$ there is a dynamical variable σ_x, called in the sequel "spin", taking values in $\{0,1\}$. With rate one each spin attempts to change its current value by tossing a coin which lands head with probability $p \in (0,1)$ and setting the new value to 1 if head and 0 if tail. However the whole operation is performed only if the current value on its right neighbor is 0. Such a model is known under the name of the East model [18] and it is easily checked to be reversible w.r.t. the product Bernoulli(p) measure. A characteristic feature of the East model is that, when $q := 1 - p \approx 0$, the relaxation to the reversible measure is extremely slow [12]:

$$T_{\text{relax}} \approx (1/q)^{\frac{1}{2}\log_2(1/q)}$$

where T_{relax} is the inverse spectral gap in the spectrum of the (self-adjoint) generator $\mathcal{L}$ of the process. Notice that if one writes $p = \frac{e^\beta}{1+e^\beta}$ then $T_{\text{relax}} \approx e^{c\beta^2}$ as $\beta \to \infty$, a behavior that is referred to as a *super-Arrhenius law* in the physics literature.

The East model is one of the simplest examples of a general class of interacting particles models which are known in physical literature as *facilitated* or *kinetically constrained spin models* (KCSM).

The common feature to all KCSM is that each dynamical variable, one for each vertex of a connected graph $\mathcal{G}$ and with values in a finite set S, waits an exponential time of mean one and then, if the surrounding current configuration satisfies a simple local constraint, is refreshed by sampling a new value from S according to some apriori specified measure ν. These models have been introduced in the physical literature [19, 20] to model the liquid/glass transition and more generally the slow "glassy" dynamics which occurs in different systems (see [32, 10] for recent review). In particular, they were devised to mimic the fact that the motion of a molecule in a dense liquid can be inhibited by the presence of too many surrounding molecules. That explains why, in all physical models, $S = \{0, 1\}$ (empty or occupied site) and the constraints specify the maximal number of particles (occupied sites) on certain sites around a given one in order to allow creation/destruction on the latter. As a consequence, the dynamics becomes increasingly slow as the density of particles, p, is increased. Moreover there usually exist *blocked configurations*, namely configurations with all creation/destruction rates identically equal to zero. This implies the existence of several invariant measures (see [26] for a somewhat detailed discussion of this issue in the context of the North-East model), the occurrence of unusually long mixing times compared to standard high-temperature stochastic Ising models and may induce the presence of ergodicity breaking transitions without any counterpart at the level of the reversible measure [17].

Because of the presence of the constraints a mathematical analysis of these models have been missing for a long time, with the notable exception of the East model [4], until a first recent breakthrough [12, 13].

In this work we partly review the results and the techniques of [12] but we also extend them in two directions. Firstly we show that the main technique can be adapted to deal with a weak interaction among the variables obtained by replacing the reversible product measure with a general high-temperature Gibbs measure. Secondly, motivated by some unpublished considerations of D. Aldous [3], we analyze

a special model, the so called *FA-1f* model, on a general connected graph and relate its relaxation time to that of the East model.

2 The Models

2.1 Setting and Notation

The models considered here are defined on a locally finite, bounded degree, connected graph $\mathcal{G} = (V, E)$ with vertex set V and edge set E. The associated graph distance will be denoted by $d(\cdot, \cdot)$ and the degree of a vertex x by Δ_x. The set of neighbors of x, *i.e.* $y \in V$ such that $d(y, x) = 1$, will be denoted by $\mathcal{N}_x$. For every subset $V' \subset V$ we denote by $\partial V'$ the set of vertices in $V \setminus V'$ with one neighbor in V'. In most cases the graph G will either be the d-dimensional lattice $\mathbb{Z}^d$ or a finite portion of it and in both cases we need some additional notation that we fix now. For any vertex $x \in \mathbb{Z}^d$ we define the $*$, the oriented and the $*$-oriented neighborhood of x as

$$\begin{aligned}
\mathcal{N}_x^* &= \{y \in \mathbb{Z}^d : \ y = x + \textstyle\sum_{i=1}^d \alpha_i \mathbf{e}_i, \ \alpha_i = \pm 1, 0 \text{ and } \sum_i \alpha_i^2 \neq 0\} \\
\mathcal{K}_x &= \{y \in \mathcal{N}_x : \ y = x + \textstyle\sum_{i=1}^d \alpha_i \mathbf{e}_i, \ \alpha_i \geq 0\} \\
\mathcal{K}_x^* &= \{y \in \mathcal{N}_x^* : \ y = x + \textstyle\sum_{i=1}^d \alpha_i \mathbf{e}_i, \ \alpha_i = 1, 0\}
\end{aligned}$$

where $\mathbf{e}_i$ are the basis vactors of $\mathbb{Z}^d$. Accordingly, the oriented and *-oriented neighborhoods $\partial_+ \Lambda$, $\partial_+^* \Lambda$ of a finite subset $\Lambda \subset \mathbb{Z}^d$ are defined as $\partial_+ \Lambda := \{\cup_{x \in \Lambda} \mathcal{K}_x\} \setminus \Lambda$, $\partial_+^* \Lambda := \{\cup_{x \in \Lambda} \mathcal{K}_x^*\} \setminus \Lambda$. A rectangle R will be a set of sites of the form

$$R := [a_1, b_1] \times \cdots \times [a_d, b_d]$$

while the collection of finite subsets of $\mathbb{Z}^d$ will be denoted by $\mathbb{F}$.

2.2 The Probability Space

Let (S, ν) be a finite probability space with $\nu(s) > 0$ for any $s \in S$. $G \subset S$ will denote a distinguished event in S, often referred to as the set of "good states", and $q \equiv \nu(G)$ its probability.

Given (S, ν) we will consider the configuration space $\Omega \equiv \Omega_V = S^V$ whose elements will be denoted by Greek letters $(\omega, \eta \dots)$. If $\mathcal{G}' = (V', E')$ is a subgraph of $\mathcal{G}$ and $\omega \in \Omega_V$ we will write $\omega_{V'}$ for its restriction to V'. We will also say that a vertex x is good for the configuration ω if $\omega_x \in G$.

On Ω equipped with the natural σ-algebra we will consider the product measure $\mu := \prod_{x\in V} \nu_x$, $\nu_x \equiv \nu$. If $\mathcal{G}' = (V', E')$ is a subgraph of $\mathcal{G}$ we will write $\mu_{V'}$ or $\mu_{\mathcal{G}'}$ for the restriction of μ to $\Omega_{V'}$. Finally, for any $f \in L^1(\mu)$, we will use the shorthand notation $\mu(f)$ to denote its expected value and $\mathrm{Var}(f)$ for its variance (when it exists).

2.3 The Markov Process

The general interacting particle models that will be studied here are Glauber type Markov processes in Ω, reversible w.r.t. the measure μ and characterized by a finite collection of influence classes $\{\mathcal{C}_x\}_{x\in V}$, where $\mathcal{C}_x$ is just a collection of subsets of V (often of the neighbors of the vertex x) satisfying the following general hypothesis:

Hp1 For all $x \in V$ and all $A \in \mathcal{C}_x$ the vertex x does not belong to A.
Hp2 $r := \sup_x \sup_{A\in\mathcal{C}_x} d(x, A) < +\infty$.

In turn the influence classes together with the good event G are the key ingredients to define the constraints of each model.

Definition 2.1. *Given a vertex $x \in V$ and a configuration ω, we will say that the constraint at x is satisfied by ω if the indicator*

$$c_x(\omega) = \begin{cases} 1 & \text{if there exists a set } A \in \mathcal{C}_x \text{ such that } \omega_y \in G \text{ for all } y \in A \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

is equal to one.

Remark 2.2. The two general hypotheses above tell us that in order to check whether the constraint is satisfied at a given vertex we do not need to check the current state of the *vertex itself* and we only need to check *locally* around the vertex. This last requirement can actually be weakened and indeed, in order to analyze certain spin exchange kinetically constrained models [11], a very efficient tool is to consider *long range* constraints!

The process that will be studied in the sequel can then be informally described as follows. Each vertex x waits an independent mean one exponential time and then, provided that the current configuration ω satisfies the constraint at x, the value ω_x is refreshed with a new value in S sampled from ν and the whole procedure starts again.

The generator $\mathcal{L}$ of the process can be constructed in a standard way (see e.g. [27, 26]) and it is a non-positive self-adjoint operator on $L^2(\Omega, \mu)$ with domain $Dom(\mathcal{L})$ and Dirichlet form given by

$$\mathcal{D}(f) = \sum_{x \in V} \mu\left(c_x \operatorname{Var}_x(f)\right), \quad f \in Dom(\mathcal{L})$$

Here $\operatorname{Var}_x(f) \equiv \int d\nu(\omega_x) f^2(\omega) - \left(\int d\nu(\omega_x) f(\omega)\right)^2$ denotes the local variance with respect to the variable ω_x computed while the other variables are held fixed. To the generator $\mathcal{L}$ we can associate the Markov semigroup $P_t := e^{t\mathcal{L}}$ with reversible invariant measure μ.

Notice that the constraints $c_x(\omega)$ are increasing functions w.r.t the partial order in Ω for which $\omega \leq \omega'$ iff $\omega'_x \in G$ whenever $\omega_x \in G$. However that does not imply in general that the process generated by $\mathcal{L}$ is attractive in the sense of Liggett [27].

Due to the fact that in general the jump rates are not bounded away from zero, irreducibility of the process is not guaranteed and the reversible measure μ is usually not the only invariant measure (typically there exist initial configurations that are blocked forever). An interesting question when $\mathcal{G}$ is infinite is therefore whether μ is ergodic/mixing for the Markov process and whether there exist other ergodic stationary measures. To this purpose it is useful to recall the following well known result (see e.g. Theorem 4.13 in [27]).

Theorem 2.3 *The following are equivalent,*

(a) $\lim_{t\to\infty} P_t f = \mu(f)$ *in* $L^2(\mu)$ *for all* $f \in L^2(\mu)$.
(b) 0 *is a simple eigenvalue for* $\mathcal{L}$.

Clearly (a) implies that $\lim_{t\to\infty} \mu\left(f P_t g\right) = \mu(f)\mu(g)$ for any $f, g \in L^2(\mu)$, *i.e.* μ is mixing.

Remark 2.4. Even if μ is mixing there will exist in general infinitely many stationary measures, *i.e.* probability measures $\tilde{\mu}$ satisfying $\tilde{\mu} P_t = \tilde{\mu}$ for all $t \geq 0$. As an example, assume c_x not identically equal to one and take an arbitrary probability measure $\tilde{\mu}$ such that $\tilde{\mu}\left(\{S \setminus G\}^V\right) = 1$. An interesting problem is therefore to classify all the stationary ergodic measures $\tilde{\mu}$ of $\{P_t\}_{t\geq 0}$, where ergodicity means that $P_t f = f$ ($\tilde{\mu}$ a.e.) for all $t \geq 0$ implies that f is constant ($\tilde{\mu}$ a.e.). As we will see later, when $\mathcal{G} = \mathbb{Z}^2$ and for a specific choice of the constraint known as the North-East model, a rather detailed answer is now available [26].

When $\mathcal{G}$ is finite connected subgraph of an infinite graph $\mathcal{G}_\infty = (V_\infty, E_\infty)$, the ergodicity issue of the resulting continuous time Markov chain can be attacked in two ways.

The first one is to analyze the chain restricted to a suitably defined ergodic component. Although such an approach is feasible and natural in some cases (see section 6 for an example), the whole analysis becomes quite cumbersome.

Another possibility, which has several technical advantages over the first one, is to unblock certain special vertices of $\mathcal{G}$ by relaxing their constraints and restore irreducibility of the chain. A natural way to do that is to imagine to extend the configuration ω, apriori defined only in V, to the vertices in $V_\infty \setminus V$ and to keep it there frozen and equal to some reference configuration τ that will be referred to as the *boundary condition*. If enough vertices in $V_\infty \setminus V$ are good for τ, then enough vertices of $\mathcal{G}$ will become unblocked and the whole chain ergodic.

More precisely we can define the finite volume constraints with boundary condition τ as

$$c^\tau_{x,V}(\omega) := c_x(\omega \cdot \tau) \tag{2.2}$$

where c_x are the constraints for $\mathcal{G}_\infty$ defined in (2.1) and $\omega \cdot \tau \in \Omega$ denotes the configuration equal to ω inside V and equal to τ in $V_\infty \setminus V$. Notice that, for any $x \in V$, the rate $c^\tau_{x,V}(\omega)$ (2.2) depends on τ only through the indicators $\{\mathbb{1}_{\tau_z \in G}\}_{z \in \mathcal{B}}$, where $\mathcal{B}$ is the *boundary set* $\mathcal{B} := (V_\infty \setminus V) \cap (\cup_{z \in V} \mathcal{C}_z)$. Therefore, instead of fixing τ, it is enough to choose a subset $\mathcal{M} \subset \mathcal{B}$, called the *good boundary set*, and define

$$c^{\mathcal{M}}_{x,V}(\omega) := c^\tau_{x,V}(\omega) \tag{2.3}$$

where τ is *any* configuration satisfying $\tau_z \in G$ for all $z \in \mathcal{M}$ and $\tau_z \notin G$ for $z \in \mathcal{B} \setminus \mathcal{M}$. We will say that a choice of $\mathcal{M}$ is *minimal* if the corresponding chain in $\mathcal{G}$ with the rates (2.3) is irreducible and it is non-irreducible for any other choice $\mathcal{M}' \subset \mathcal{M}$. The choice $\mathcal{M} = \mathcal{B}$ will be called *maximal*. For convenience we will write $\mathcal{L}^{\max}_\Lambda$ ($\mathcal{L}^{\min}_\Lambda$) for the corresponding generators.

Remark 2.5. Without any other specification for the influence classes of the model it may very well be the case that there exists no boundary conditions for which the chain is irreducible and/or their existence may depend on the choice of the finite subgraph $\mathcal{G}$. However, as we will see later, for all the interesting models discussed in the literature all these issues will have a rather simple solution.

We will now describe some of the basic models and solve the problem of boundary conditions for each one of them.

2.4 0-1 Kinetically Constrained Spin Models

In most models considered in the physical literature the finite probability space (S, ν) is a simple $\{0,1\}$ Bernoulli space and the good set G is conventionally chosen as the empty (vacant) state $\{0\}$. Any model with these features will be called in the sequel a "0-1 KCSM" (kinetically constrained spin model). Although in most cases the underlying graph $\mathcal{G}$ is a regular lattice like $\mathbb{Z}^d$, whenever is possible we will try to work in full generality.

Given a 0-1 KCSM, the parameter $q = \nu(0)$ can be varied in $[0,1]$ while keeping fixed the basic structure of the model (*i.e.* the notion of the good set and the constraints c_x) and it is natural to define a critical value q_c as

$$q_c = \inf\{q \in [0,1] : 0 \text{ is a simple eigenvalue of } \mathcal{L}\}$$

As we will prove below q_c coincides with the *bootstrap percolation threshold* q_{bp} of the model defined as follows [34][1]. For any $\eta \in \Omega$ define the bootstrap map $T : \Omega \mapsto \Omega$ as

$$(T\eta)_x = 0 \quad \text{if either} \quad \eta_x = 0 \quad \text{or} \quad c_x(\eta) = 1. \tag{2.4}$$

Denote by $\mu^{(n)}$ the probability measure on Ω obtained by iterating n-times the above mapping starting from μ. As $n \to \infty$ $\mu^{(n)}$ converges to a limiting measure $\mu^{(\infty)}$ [34] and it is natural to define the critical value q_{bp} as

$$q_{bp} = \inf\{q \in [0,1] : \mu^{\infty} = \delta_0\}$$

where δ_0 is the probability measure assigning unit mass to the constant configuration identically equal to zero. In other words q_{bp} is the infimum of the values q such that, with probability one, the graph $\mathcal{G}$ can be entirely emptied. Using the fact that the c_x's are increasing function of η it is easy to check that $\mu^{(\infty)} = \delta_0$ for any $q > q_{bp}$.

Proposition 2.6 ([12]) *$q_c = q_{bp}$ and for any $q > q_c$ 0 is a simple eigenvalue for $\mathcal{L}$.*

Remark 2.7. In [12] the proposition has been proved in the special case $\mathcal{G} = \mathbb{Z}^d$ but actually the same arguments apply to any bounded degree connected graph.

Having defined the bootstrap percolation it is natural to divide the 0-1 KCSM into two distinct classes.

[1] In most of the bootstrap percolation literature the role of the 0's and the 1's is inverted

Definition 2.8. *We will say that a 0-1 KCSM is* non cooperative *if there exists a finite set $\mathcal{B} \subset V$ such that any configuration η which is empty in all the sites of $\mathcal{B}$ reaches the empty configuration (all 0's) under iteration of the bootstrap mapping. Otherwise the model will be called* cooperative.

Remark 2.9. Notice that for a non-cooperative model the critical value q_c is obviously zero since with μ-probability one a configuration will contain the required finite set $\mathcal{B}$ of zeros.

We will now illustrate some of the most studied models.

[1] Frederickson-Andersen (FA-jf) facilitated models [19, 20]. In the facilitated models the constraint at x requires that at least $j \leq \Delta_x$ neighbors are vacant. More formally

$$\mathcal{C}_x = \{A \subset \mathcal{N}_x : |A| \geq j\}$$

When $j = 1$ the model is non-cooperative for any connected graph $\mathcal{G}$ and ergodicity of the Markov chain is clearly guaranteed by the presence of at least one unblocked vertex. When $j > 1$ ergodicity on a general graph is more delicate and we restrict ourselves to finite rectangles R in $\mathbb{Z}^d$. In that case and for the most constrained cooperative case $j = d$ among the irreducible ones, irreducibility is guaranteed if we assume a boundary configuration identically empty on $\partial_+ R$. Quite remarkably, using results from bootstrap percolation [34] combined with proposition 2.6, when $\mathcal{G} = \mathbb{Z}^d$ and $2 \leq j \leq d$ the ergodicity threshold q_c always vanishes.

[2] Spiral model [9, 8]. This model is defined on $\mathbb{Z}^2$ with the following choice for the influence classes

$$\mathcal{C}_x = \{NE_x \cup SE_x;\ SE_x \cup SW_x;\ SW_x \cup NW_x;\ NW_x \cup NE_x\}$$

where $NE_x = (x + \mathbf{e}_2, x + \mathbf{e}_1 + \mathbf{e}_2)$, $SE_x = (x + \mathbf{e}_1, x + \mathbf{e}_1 - \mathbf{e}_2)$, $SW_x = (x - \mathbf{e}_2, x - \mathbf{e}_2 - \mathbf{e}_1)$ and $NW_x = (x - \mathbf{e}_1; x - \mathbf{e}_1 + \mathbf{e}_2)$. In other words the vertex x can flip iff either its North-East (NE_x) or its South-West (SW_x) neighbours (or both of them) are empty *and* either its North-West (NW_x) or its South-East (SE_x) neighbours (or both of them) are empty too. The model is clearly cooperative and in [8] it has been proven that its critical point q_c coincides with $1 - p_c^o$, where p_c^o is the critical threshold for oriented percolation. The interest of this model lies on the fact that its bootstrap percolation is expected to

display a peculiar mixed discontinuous/critical character which makes it relevant as a model for the liquid glass and more general jamming transitions [9, 8].

[3] Oriented models. Oriented models are similar to the facilitated models but the neighbors of a given vertex x that must be vacant in order for x to become free to flip, are chosen according to some orientation of the graph. Instead of trying to describe a very general setting we present three important examples.

Example 2.10. The first and best known example is the so called East model [18]. Here $\mathcal{G} = \mathbb{Z}$ and for every $x \in \mathbb{Z}$ the influence class $\mathcal{C}_x$ consists of the vertex $x + 1$. In other words any vertex can flip iff its right neighbor is empty. The minimal boundary conditions in a finite interval which ensure irreducibility of the chain are of course empty right boundary, *i.e.* the rightmost vertex is always unconstrained. The model is clearly cooperative but $q_c = 0$ since in order to empty $\mathbb{Z}$ it is enough to start from a configuration for which any site x has some empty vertex to its right. One could easily generalize the model to the case when $\mathcal{G}$ is a rooted tree (see section 6). In that case any vertex different from the root can be updated iff its ancestor is empty. The root itself is unconstrained.

Example 2.11. The second example is the North-East model in $\mathbb{Z}^2$ [25]. Here one chooses $\mathcal{C}_x$ as the North and East neighbor of x. The model is clearly cooperative and its critical point q_c coincides with $1 - p_c^o$, where p_c^o is the critical threshold for oriented percolation in $\mathbb{Z}^2$ [34]. For such a model much more can be said about the stationary ergodic measures of the Markov semigroup P_t.

Theorem 2.12 ([26]) *If $q < q_c$ the trivial measure δ_1 that assigns unit mass to the configuration identically equal to 1 is the only translation invariant, ergodic, stationary measure for the system. If $q \geq q_c$ the reversible measure μ is the unique, non trivial, ergodic, translation invariant, stationary measure.*

Example 2.13. The third model was suggested in [4] and it is defined on a rooted (finite or infinite) binary tree $\mathcal{T}$. Here a vertex x can flip iff its two children are vacant. If the tree is finite then ergodicity requires that all the leaves of $\mathcal{T}$ are unconstrained. It is easy to check that the critical threshold satisfies $q_c = 1/2$, the site percolation threshold on the binary tree.

3 Quantities of Interest and Related Problems

Back to the general model we now define two main quantities that are of mathematical and physical interest.

The first one is the spectral gap of the generator $\mathcal{L}$, defined as

$$\mathrm{gap}(\mathcal{L}) := \inf_{f \neq \mathrm{const}} \frac{\mathcal{D}(f)}{\mathrm{Var}(f)} \tag{3.1}$$

A positive spectral gap implies that the reversible measure μ is mixing for the semigroup P_t with exponentially decaying correlations:

$$\mathrm{Var}\,(P_t f) \leq e^{-2t\,\mathrm{gap}(\mathcal{L})}\,\mathrm{Var}(f), \qquad \forall\, f \in L^2(\mu).$$

Remark 3.1. In the sequel the time scale $T_{\mathrm{rel}} := \mathrm{gap}^{-1}$ which is naturally fixed by the spectral gap will be referred to as the *relaxation time* of the process.

For a 0-1 KCSM, two natural questions arise.

1. Define the new critical point $q_c' := \inf\{q \in [0,1] :\ \mathrm{gap}(\mathcal{L}) > 0\}$. Obviously $q_c' \geq q_c$. Is it the case that equality holds?
2. If $q_c' = q_c$ what is the behaviour of $\mathrm{gap}(\mathcal{L})$ as $q \downarrow q_c$?

As we will see later for most of the relevant models it is possible to answer in rather detailed way to both questions.

The second quantity of interest is the so called persistence function (see e.g. [23, 36]) defined by

$$F(t) := \int d\mu(\eta)\ \mathbb{P}(\sigma_0^\eta(s) = \eta_0,\ \forall s \leq t) \tag{3.2}$$

where $\{\sigma_s^\eta\}_{s\geq 0}$ denotes the process started from the configuration η. In some sense the persistence function, a more accessible quantity to numerical simulation than the spectral gap, provides a measure of the "mobility" of the system. Here the main questions are:

1. What is the behavior of $F(t)$ for large time scales?
2. For a 0-1 KCSM is it the case that $F(t)$ decays exponentially fast as $t \to \infty$ for any $q > q_c'$?
3. If the answer to the previous question is positive, is the decay rate related to the spectral gap in a simple way or the decay rate of $F(t)$ requires a deeper knowledge of the spectral density of $\mathcal{L}$?
4. Is it possible to exhibit examples of 0-1 KCSM in which the persistence function shows a crossover between a stretched and a pure exponential decay?

Unfortunately the above questions are still mostly unanswered except for the first two.

3.1 Some Useful Observations to Bound the Spectral Gap

It is important to observe the following kind of monotonicity that can be exploited in order to bound the spectral gap of one model with the spectral gap of another one.

Definition 3.2. *Suppose that we are given two influence classes $\mathcal{C}_0$ and $\mathcal{C}'_0$, denote by $c_x(\omega)$ and $c'_x(\omega)$ the corresponding rates and by $\mathcal{L}$ and $\mathcal{L}'$ the associated generators on $L^2(\mu)$. If, for all $\omega \in \Omega$ and all $x \in V$, $c'_x(\omega) \leq c_x(\omega)$, we say that $\mathcal{L}$ is dominated by $\mathcal{L}'$.*

Remark 3.3. The term domination here has the same meaning it has in the context of bootstrap percolation. It means that the KCSM associated to $\mathcal{L}'$ is *more* constrained than the one associated to $\mathcal{L}$.

Clearly, if $\mathcal{L}$ is dominated by $\mathcal{L}'$, $\mathcal{D}'(f) \leq \mathcal{D}(f)$ and therefore $\mathrm{gap}(\mathcal{L}') \leq \mathrm{gap}(\mathcal{L})$.

Example 3.4. Assume that the graph $\mathcal{G}$ has n vertices and contains a Hamilton path $\Gamma = \{x_1, x_2, \ldots, x_n\}$, *i.e.* $d(x_{i+1}, x_i) = 1$ for all $1 \leq i \leq n-1$ and $x_i \neq x_j$ for all $i \neq j$. Consider the FA-1f model on $\mathcal{G}$ with one special vertex, e.g. x_n, unconstrained ($c_{x_n} \equiv 1$). Then, if we replace $\mathcal{G}$ by Γ equipped with its natural graph structure and we denote by $\mathcal{L}$ and $\mathcal{L}'$ the respective generators, we get that $\mathrm{gap}(\mathcal{L}) \geq \mathrm{gap}(\mathcal{L}')$. Clearly $\mathcal{L}'$ describes the FA-1f model on the finite interval $[1, \ldots, n] \subset \mathbb{Z}$ with the last vertex free to flip. This in turn is dominated by $\mathcal{L}_{\mathrm{East}}$, the generator of the East model on $[1, \ldots, n]$, which is known to have a positive [4, 12] spectral gap uniformly in n. Therefore the latter result holds also for $\mathrm{gap}(\mathcal{L}')$ and $\mathrm{gap}(\mathcal{L})$.

Example 3.5. Along the lines of the previous example we could lower bound the spectral gap of the FA-2f model in $\mathbb{Z}^d$, $d \geq 2$, with that in $\mathbb{Z}^2$, by restricting the sets $A \in \mathcal{C}_0$ to e.g. the $(\mathbf{e}_1, \mathbf{e}_2)$-plane.

For a last and more detailed example of the comparison technique we refer the reader to section 6.

Although the comparison technique can be quite effective in proving positivity of the spectral gap, one should keep in mind that, in general, it provides quite poor bounds, particularly in the limiting case $q \downarrow q_c$.

The second observation we make consists in relating $\mathrm{gap}(\mathcal{L})$ when the underlying graph is infinite to its finite graph analogue. Fix $r \in V$ and let $\mathcal{G}_{n,r} \subset \mathcal{G}$ be the connected ball centered at r of radius n. Suppose that $\inf_n \mathrm{gap}(\mathcal{L}^{\max}_{\mathcal{G}_{n,r}}) > 0$. It is then easy to conclude that $\mathrm{gap}(\mathcal{L}) > 0$.

Indeed, following Liggett Ch.4 [27], for any $f \in Dom(\mathcal{L})$ with $\mathrm{Var}(f) > 0$ pick $f_n \in L^2(\Omega, \mu)$ depending only on finitely many spins

so that $f_n \to f$ and $\mathcal{L}f_n \to \mathcal{L}f$ in L^2. Then $\mathrm{Var}(f_n) \to \mathrm{Var}(f)$ and $\mathcal{D}(f_n) \to \mathcal{D}(f)$. But since f_n depends on finitely many spins

$$\mathrm{Var}(f_n) = \mathrm{Var}_{\mathcal{G}_{m,r}}(f_n) \quad \text{and} \quad \mathcal{D}(f_n) = \mathcal{D}_{\mathcal{G}_{m,r}}(f_n)$$

provided that m is a large enough square (depending on f_n). Therefore

$$\frac{\mathcal{D}(f)}{\mathrm{Var}(f)} \geq \inf_n \mathrm{gap}(\mathcal{L}_{\mathcal{G}_{n,r}}) > 0.$$

and $\mathrm{gap}(\mathcal{L}) \geq \inf_n \mathrm{gap}(\mathcal{L}_{\mathcal{G}_{n,r}}) > 0$.

4 Main Results for 0-1 KCSM on Regular Lattices

In this section we state some of the main results for a general 0-1 KCSM on $\mathbb{Z}^d$ which have been obtained in [12].

Fix an integer length scale ℓ larger than the range of the constraints and let $\mathbb{Z}^d(\ell) \equiv \ell\,\mathbb{Z}^d$. Consider a partition of $\mathbb{Z}^d$ into disjoint rectangles $\Lambda_z := \Lambda_0 + z$, $z \in \mathbb{Z}^d(\ell)$, where $\Lambda_0 = \{x \in \mathbb{Z}^d : \ 0 \leq x_i \leq \ell - 1, \ i = 1, .., d\}$.

Definition 4.1. *Given $\epsilon \in (0,1)$ we say that $G_\ell \subset \{0,1\}^{\Lambda_0}$ is a ϵ-good set of configurations on scale ℓ if the following two conditions are satisfied:*

(a) $\mu(G_\ell) \geq 1 - \epsilon$.

(b) For any collection $\{\xi^{(x)}\}_{x \in \mathcal{K}_0^*}$ of spin configurations such that $\xi^{(x)} \in G_\ell$ for all $x \in \mathcal{K}_0^*$, the following holds. For any $\xi \in \Omega$ which coincides with $\xi^{(x)}$ in $\cup_{x \in \mathcal{K}_0^*} \Lambda_{\ell x}$, there exists a sequence of legal moves inside $\cup_{x \in \mathcal{K}_0^*} \Lambda_{\ell x}$ (i.e. single spin moves compatible with the constraints) which transforms ξ into a new configuration $\tau \in \Omega$ such that the Markov chain in Λ_0 with boundary conditions τ is ergodic.

Remark 4.2. In general the transformed configuration τ will be identically equal to zero on $\partial_+^* \Lambda_0$. It is also clear that assumption (b) has been made having in mind models like the East, the FA-jf or the N-E which, modulo rotations, are dominated by a model with influence class $\tilde{\mathcal{C}}_x$ *entirely contained* in the sector $\{y : y = x + \sum_{i=1}^d \alpha_i \mathbf{e}_i, \ \alpha_i \geq 0\}$. If this is not the case one should instead use a non rectangular geometry for the tiles of the partition of $\mathbb{Z}^d$, adapted to the choice of the influence classes. For example for the Spiral Model the basic tile at length scale ℓ is a quadrangular region $\mathcal{R}_0$ with one side parallel to $\mathbf{e}_1$ and two sides parallel to $\mathbf{e}_1 + \mathbf{e}_2$, $\mathcal{R}_0 := \cup_1^\ell S_0 + (i-1)(\mathbf{e}_1 + \mathbf{e}_2)$

with $S_0 := \{x \in \mathbb{Z}^2 : 0 \le x_1 \le \ell - 1, x_2 = 0\}$. In this case condition (b) should also be modified by substituting everywhere $\partial^*_+ \Lambda_0$ with $\widetilde{\partial}^*_+ \Lambda_0 := \mathbf{e}_1, \mathbf{e}_1 - \mathbf{e}_2, -e_2$.

With the above notation the first main result of [12] can be formulated as follows.

Theorem 4.3 *There exists a universal constant* $\epsilon_0 \in (0,1)$ *such that, if there exists* ℓ *and a* ϵ_0*-good set* G_ℓ *on scale* ℓ*, then* $\inf_{\Lambda \in \mathbb{F}} \mathrm{gap}(\mathcal{L}^{\max}_\Lambda) > 0$*. In particular* $\mathrm{gap}(\mathcal{L}) > 0$*.*

In several examples, e.g. the FA-jf models, the natural candidate for the event G_ℓ is the event that the tile Λ_0 is "internally spanned", a notion borrowed from bootstrap percolation [2, 34, 14, 24, 15]:

Definition 4.4. *We say that a finite set* $\Gamma \subset \mathbb{Z}^d$ *is* internally spanned *by a configuration* $\eta \in \Omega$ *if, starting from the configuration* η^Γ *equal to one outside* Γ *and equal to* η *inside* Γ*, there exists a sequence of legal moves inside* Γ *which connects* η^Γ *to the configuration identically equal to zero inside* Γ *and identically equal to one outside* Γ*.*

Of course whether or not the set Λ_0 is internally spanned for η depends only on the restriction of η to Λ_0. One of the major results in bootstrap percolation problems has been the exact evaluation of the μ-probability that the box Λ_0 is internally spanned as a function of the length scale ℓ and the parameter q [24, 34, 14, 15, 2]. For non-cooperative models it is obvious that for any $q > 0$ such probability tends very rapidly (exponentially fast) to one as $\ell \to \infty$, since the existence of at least one completely empty finite set $\mathcal{B}+x \subset \Lambda_0$ (see definition 2.8), allows to empty all Λ_0. For some cooperative systems like e.g. the FA-2f in $\mathbb{Z}^2$, it has been shown that for any $q > 0$ such probability tends very rapidly (exponentially fast) to one as $\ell \to \infty$ and that it abruptly jumps from being very small to being close to one as ℓ crosses a critical scale $\ell_c(q)$. In most cases the critical length $\ell_c(q)$ diverges very rapidly as $q \downarrow 0$. Therefore, for such models and $\ell > \ell_c(q)$, one could safely take G_ℓ as the collection of configurations η such that Λ_0 is internally spanned for η. We now formalize what we just said.

Corollary 4.5 *Assume that* $\lim_{\ell \to \infty} \mu(\Lambda_0$ *is internally spanned* $) = 1$ *and that the Markov chain in* Λ_0 *with zero boundary conditions on* $\cup_{x \in \mathcal{K}^*_0} \Lambda_{\ell x}$ *is ergodic. Then* $\mathrm{gap}(\mathcal{L}) > 0$*.*

We stress that for some models a notion of good event which differs from requiring internal spanning is needed. This is the case for the

N-E and Spiral models, as can be immediately seen by noticing that at any length scale it is possible to construct small clusters of particles in proper corners of the tiles that can never be erased by internal moves. The choice of the proper ϵ-good set of configurations for N-E has already been discussed in [12]. For the Spiral Model the definition which naturally arises from the results in [9] is the following. Let $\widetilde{\mathcal{R}}_0$ be the region obtained from $\mathcal{R}_0$ by subtracting two proper quadrangular regions at the bottom left and top right corners, namely $\widetilde{\mathcal{R}}_0 := \mathcal{R}_0 \setminus (\mathcal{R}_{bl} \cup \mathcal{R}_{tr})$ where $\mathcal{R}_{bl}$ ($\mathcal{R}_{tr}$) have the same shape of $\mathcal{R}_0$ shrinked at length scale $\ell/4$ and have the bottom left (top right) corner which coincides with the one of $\mathcal{R}_0$. The ϵ-good set of configurations on scale ℓ, G_ℓ, includes all configurations η such that there exists a sequence of legal moves inside $\mathcal{R}_0$ which connects $\eta^{\mathcal{R}_0}$ (the configuration which has all ones outside $\mathcal{R}_0$ and equals η inside) to a configuration identically equal to zero inside $\widetilde{\mathcal{R}}_0$. Lemma 4.7 and Proposition 4.9 of [9] prove, respectively, property (a) and (b) of Definition 4.1 (with $\partial^*_+ \Lambda_0$ substituted with $\mathbf{e}_1, \mathbf{e}_1 - \mathbf{e}_2, -e_2$, see remark 4.2) when the density is below the critical density of oriented percolation. Thus, using this definition for the good event and Theorem 4.3 we conclude that

Theorem 4.6 $\mathrm{gap}(\mathcal{L}_{spiral}) > 0$ *at any* $\rho < p_c^o$.

The second main result concerns the long time behavior of the persistence function $F(t)$ defined in (3.2).

Theorem 4.7 *Assume that* $\mathrm{gap}(\mathcal{L}) > 0$. *Then* $F(t) \le e^{-q\,\mathrm{gap}\,t} + e^{-p\,\mathrm{gap}\,t}$.

Remark 4.8. The above theorems disprove some conjectures which appeared in the physics literature [21, 23, 5, 6], based on numerical simulations and approximate analytical treatments, on the existence of a second critical point $q_c' > q_c$ at which the spectral gap vanishes and/or below which $F(t)$ would decay in a stretched exponential form $\simeq \exp(-t/\tau)^\beta$ with $\beta < 1$.

Theorem 4.7 also indicates that one can obtain *upper bounds* on the spectral gap by proving *lower bounds* on the persistence function. Concretely a lower bound on the persistence function can be obtained by restricting the μ-average to those initial configurations η for which the origin is blocked with high probability for all times $s \le t$. Unfortunately in most models such a strategy leads to lower bound on $F(t)$ which are usually quite far from the above upper bound and it is an interesting open problem to find an exact asymptotic as $t \to \infty$ of $F(t)$.

Finally we observe that for the North-East model on $\mathbb{Z}^2$ at the critical value $q = q_c$ the spectral gap vanishes and the persistence function satisfies $\int_0^\infty dt\, F(\sqrt{t}) = \infty$ (see Theorem 6.17 and Corollary 6.18 in [12]).

4.1 Some Ideas of the Strategy for Proving Theorems 4.3, 4.7

The main idea behind the proof of theorem 4.3 goes as follows. First of all one covers the lattice with non overlapping cubic blocks $\{\Lambda_{\ell x}\}_{x\in\mathbb{Z}^d}$ and, on the rescaled lattice $\mathbb{Z}^d(\ell) := \ell\mathbb{Z}^d$, one considers the new model with single spin space $S = \{0,1\}^{\ell^d}$, good event $G := G_\ell$, single site measure the restriction of μ to S and renormalized constraints $\{c_x^{ren}\}_{x\in\mathbb{Z}^d(\ell)}$ which are a strengthening of the North-East ones namely

$$c_x^{ren}(\eta) = 1 \quad \text{iff } \eta_y \in G \text{ for all } y \in \mathcal{K}_x^*.$$

Such a model is referred to in [12] as the **-general model*. By assumption the probability of G can be made arbitrarily close to one by taking ℓ large enough and therefore, by the so called *Bisection-Constrained approach* which is detailed in the next section for the case when μ is a high temperature Gibbs measure, the spectral gap of the *-general model is positive. Next one observes that assumption (b) of the theorem is there exactly to allow one to reconstruct any *legal* move of the *-general model, *i.e.* a full update of an entire block of spins, by means of a finite (depending only on ℓ) sequence of *legal* moves for the original 0-1 KCMS. It is then an easy step, using standard path techniques for comparing two different Markov chains (see e.g. [33]), to go from the Poincaré inequality for the *-general model to the Poincaré inequality for the original model.

The proof of (a slightly less precise version of) Theorem 4.7 given in [12] is based on the Feynman-Kac formula and standard large deviation considerations. However it is possible to provide a simpler and more precise argument as follows. One first observe that $F(t) = F_1(t) + F_0(t)$ where

$$F_1(t) = \int d\mu(\eta)\, \mathbb{P}(\sigma_0^\eta(s) = 1 \text{ for all } s \le t)$$

and similarly for $F_0(t)$. Consider now $F_1(t)$, the case of $F_0(t)$ being similar, and define $T_A(\eta)$ as the hitting time of the set $A := \{\eta :\ \eta_0 = 0\}$ starting from the configuration η. Then (see e.g. Theorem 2 in [1])

$$F_1(t) = \mathbb{P}_\mu\Big(T_A > t\Big) \le e^{-t\lambda_A}$$

where $\mathbb{P}_\mu$ denotes the probability over the process started from the equilibrium distribution μ and λ_A is given by the variational formula for the Dirichlet problem

$$\lambda_A := \inf\Big\{\mathcal{D}(f) : \ \mu(f^2) = 1, \ f \equiv 0 \text{ on } A\Big\} \tag{4.1}$$

Notice that for any f as above $\mathrm{Var}(f) \geq \mu(A) = q$. Therefore $\lambda_A \geq q$ gap and the proof is complete.

4.2 Asymptotics of the Spectral Gap Near the Ergodicity Threshold

An important question, particularly in connection with numerical simulations or non-rigorous approaches, is the behavior near the ergodicity threshold q_c of the spectral gap for each specific model. Here is a set of results proven in [12].

1. **East Model.**

$$\lim_{q\to 0} \log(1/\,\mathrm{gap})/(\log(1/q))^2 = (2\log 2)^{-1} \tag{4.2}$$

2. **FA-1f.** For any $d \geq 1$, there exists a constant $C = C(d)$ such that for any $q \in (0,1)$, the spectral gap on $\mathbb{Z}^d$ satisfies:

$$\begin{aligned} C^{-1}q^3 \leq \mathrm{gap}(\mathcal{L}) \leq Cq^3 \qquad & \text{for } d = 1, \\ C^{-1}q^2/\log(1/q) \leq \mathrm{gap}(\mathcal{L}) \leq Cq^2 \qquad & \text{for } d = 2, \\ C^{-1}q^2 \leq \mathrm{gap}(\mathcal{L}) \leq Cq^{1+\frac{2}{d}} \qquad & \text{for } d \geq 3. \end{aligned}$$

3. **FA-df in $\mathbb{Z}^d$.** Fix $\epsilon > 0$. Then there exists $c = c(d)$ such that

$$\begin{aligned} \Big[\exp^{d-1}(c/q^2)\Big]^{-1} \leq \mathrm{gap}(\mathcal{L}) \leq \Big[\exp^{d-1}\big(\frac{\lambda_1 - \epsilon}{q}\big)\Big]^{-1} \qquad & d \geq 3 \\ \exp(-c/q^5) \leq \mathrm{gap}(\mathcal{L}) \leq \exp\big(-\frac{(\lambda_1 - \epsilon)}{q}\big) \qquad & d = 2 \end{aligned} \tag{4.3}$$

   as $q \downarrow 0$, where $\exp^{d-1}$ denote the $(d-1)^{\text{th}}$-iterate of the exponential function and $\lambda_1 = \pi^2/18$.

The proof of the lower bounds is a rather delicate combination of the renormalization scheme described above together with paths techniques as described in [33]. The upper bounds are proved instead either by a careful choice of a test function in the variational characterization of the spectral gap or by a lower bound on the persistence function $F(t)$ combined with the upper bound given in Theorem 4.7.

Remark 4.9. Again some of the above findings disprove previous claims for the East model [36] and for the FA-1f model in $d = 2, 3$ [6].

The result for the East model actually came out as a surprise. In [36] the model was considered "essentially" solved and the result for the spectral gap was gap $\approx q^{\log_2(q)}$ as $q \downarrow 0$ to be compared to the correct scaling $q^{\log_2(q)/2}$. In [4] the above solution was proved to be a *lower* bound and an upper bound of the form $q^{\log_2(q)/2}$ was rigorously established but considered poor because off by a power 1/2 from the supposedly correct behavior.

The scaling indicated in [36] is based in part on the following consideration. Fix $q \ll 1$ and consider the East model on the interval $\Lambda_q := [0, \dots, 1/q]$ with the last site free to flip (*i.e.* zero boundary conditions). Notice that $1/q$ is the average distance between the zeros. Start from the configuration identically equal to one and let T be the (random) time at which the origin is able to flip. Energy barriers consideration (see [3, 4, 16]) suggest that $\mathbb{E}(T)$ should scale as $q^{\log_2(q)}$ and that is what was assumed in [36]. However it is not difficult to prove that the scaling of $\mathbb{E}(T)$ is bounded above by $(q\,\mathrm{gap})^{-1}$. Indeed we can write for any $t \geq 0$

$$\exp(-cq\,\mathrm{gap}(\mathcal{L}_{\Lambda_q})t) \geq \tilde{F}(t) \geq \mu(\text{all ones})\mathbb{P}(T \geq t) \geq e^{-2}\mathbb{P}(T \geq t)$$

where $\tilde{F}(t)$ is the finite volume persistence function. Integrating over t and using the monoticity of the gap (see [12, Lemma 2.11]) give $\mathbb{E}(T) \leq e^2 c(q\,\mathrm{gap}(\mathcal{L}_{\Lambda_q})))^{-1} \leq e^2 c(q\,\mathrm{gap}(\mathcal{L}))^{-1}$. This, in view of Theorem 4.3, is incompatible with the assumed scaling $q^{\log_2(q)}$.

Moreover one can obtain a lower bound on $\mathbb{E}(T)$ as follows. Let λ be such that $\mathbb{P}(T \geq \lambda) = e^{-1}$ then clearly $\mathbb{P}(T \geq t) \leq e^{-\lfloor t/\lambda \rfloor}$ and $\mathbb{E}(T) \geq e^{-1}\lambda$. We can always couple in the natural way two copies of the process, one started from all ones and the other from any other configuration η, and conclude that

$$\mathbb{P}(\text{the two copies have not coupled at time } t) \leq \mathbb{P}(T \geq t) \leq e^{1-\lambda t}.$$

Standard arguments give immediately that $\mathrm{gap}^{-1} \leq \lambda$ *i.e.* $\mathbb{E}(T) \geq e^{-1}\,\mathrm{gap}^{-1}$. In conclusion

$$e^{-1}\big[\mathrm{gap}(\mathcal{L}_{\Lambda_q})\big]^{-1} \leq \mathbb{E}(T) \leq e^2 c\big[q\,\mathrm{gap}(\mathcal{L}_{\Lambda_q})\big]^{-1}$$

5 Extension to Interacting Models

In this section we show how to extend the results on the positivity of the spectral gap for 0-1 KCSM on a regular lattice $\mathbb{Z}^d$ to the case in which a weak interaction is present among the spins. We begin by defining what we mean by an *interaction*.

Definition 5.1. *A finite range interaction Φ is a collection $\Phi := \{\Phi_\Lambda\}_{\Lambda\in\mathbb{F}}$ where*

i) $\Phi_\Lambda : \Omega_\Lambda \mapsto \mathbb{R}$ *for every* $\Lambda \in \mathbb{F}$*;*
ii) $\Phi_\Lambda = 0$ *if* $\mathrm{diam}(\Lambda) \geq r$ *for some finite* $r = r(\Phi)$ *called the range of the interaction;*
iii) $\|\Phi\| \equiv \sup_{x\in\mathbb{Z}^d} \sum_{\Lambda\ni x} \|\Phi_\Lambda\|_\infty < \infty$*;*

We will say that $\Phi \in \mathcal{B}_{M,r}$ *if* $r(\Phi) \leq r$ *and* $\|\Phi\| \leq M$*.*

Given an interaction $\Phi \in \mathcal{B}_{r,M}$ and $\Lambda \in \mathbb{F}$, we define the energy in Λ of a spin configuration $\sigma \in \Omega$ by

$$H_\Lambda(\sigma) = \sum_{A\cap\Lambda\neq\emptyset} \Phi_A(\sigma)$$

For $\sigma \in \Omega_\Lambda$ and $\tau \in \Omega_{\Lambda^c}$ we also let $H^\tau_\Lambda(\sigma) := H_\Lambda(\sigma\cdot\tau)$ where $\sigma\cdot\tau$ denotes the configuration equal to σ inside Λ and to τ outside it. Finally, for any $\Lambda \in \mathbb{F}$ and $\tau \in \Omega_{\Lambda^c}$, we define the finite volume Gibbs measure on Ω_Λ with boundary conditions τ and apriori single spin measure ν by the formula

$$\mu^{\Phi,\tau}_\Lambda(\sigma) := \frac{1}{Z^{\Phi,\tau}_\Lambda} e^{-H^\tau_\Lambda(\sigma)} \prod_{x\in\Lambda} \nu(\sigma_x)$$

where $Z^{\Phi,\tau}_\Lambda$ is a normalization constant.

The key property of Gibbs measures is that, for any $V \subset \Lambda$ and any ξ in $\Lambda\backslash V$, the conditional Gibbs measure in Λ with boundary conditions τ given ξ coincides with the Gibbs measure in V with boundary condition $\tau_{\Lambda^c}\cdot\xi$. More formally

$$\mu^{\Phi,\tau}_\Lambda(\cdot \,|\, \sigma_{V^c} = \xi) = \mu^{\Phi,\tau_{\Lambda^c}\cdot\xi}_V(\cdot)$$

Clearly averages w.r.t. $\mu^{\Phi,\tau}_\Lambda(\cdot \,|\, \sigma_{V^c} = \xi)$ are function of ξ and, whenever confusion does not arise, we will systematically drop ξ from our notation.

As it is well known (see e.g. [35]), for any $r < \infty$ there exists $M_0 > 0$ such that for any $0 < M < M_0$ the following holds. For any $\Phi \in \mathcal{B}_{r,M}$ there exists a unique probability measure μ^Φ on Ω, called the unique Gibbs measure associated to the interaction Φ with apriori measure ν, such that, for any τ,

$$\lim_{\Lambda\uparrow\mathbb{Z}^d} \mu^{\Phi,\tau}_\Lambda = \mu^\Phi$$

where the limit is to be understood as a weak limit. Moreover the limit is reached "exponentially fast" in the strongest possible sense. Namely, for any $\Delta \subset \Lambda \in \mathbb{F}$ and any two boundary conditions τ, τ',

$$\max_{\sigma_\Delta} \left| \frac{\mu_\Lambda^{\Phi,\tau'}(\sigma_\Delta)}{\mu_\Lambda^{\Phi,\tau}(\sigma_\Delta)} - 1 \right| \le K |D_r(\tau,\tau')|\, e^{-md\left(\Delta,\, D_r(\tau,\tau')\right)} \tag{5.1}$$

where $D_r(\tau,\tau') = \{y :\ 0 < d(y,\Lambda) \le r,\ \tau_y \ne \tau'_y\}$ and the constants m, K depend only on M, r, d. Moreover $m \uparrow +\infty$ as $M \downarrow 0$. When $d = 1$ the threshold M_0 can be taken equal to $+\infty$. In all what follows we will always assume that $\Phi \in \mathcal{B}_{r,M}$ for some apriori given r, M and that $M < M_0$.

Remark 5.2. In general the constant M_0 does not coincide with any "critical point" for the model. It is only a sort of "high temperature threshold" (see [29, 30, 31] for more details about this issue).

Having described the notion of the unique Gibbs measure corresponding to Φ, we can define the generator $\mathcal{L}^\Phi$ of a 0-1 KCSM with interaction Φ and constraints c_x given by (2.1), as the unique self-adjoint operator on $L^2(\Omega, \mu^\Phi)$ with quadratic form

$$\mathcal{D}^\Phi(f) = \sum_x \mu^\Phi \left(c_x \operatorname{Var}_x^\Phi(f) \right), \qquad f \text{ local}$$

where now the local variance $\operatorname{Var}_x^\Phi(f)$ is computed with the conditional Gibbs measure given all the spins outside x. The construction of the generator in a finite volume Λ with boundary conditions τ is exactly the same as in the non-interacting case and we skip it.

5.1 Spectral Gap for a Weakly Interacting North-East Model

Instead of trying to prove a very general result on the spectral gap of a weakly interacting KCSM, we will explain how to deal with the interaction in the concrete case of the North-East model introduced in section 2.4. Moreover, in order not to obscure the discussion with renormalization or block constructions, we will make the unnecessary assumption that the basic parameter q of the reference measure ν is very close to one.

Theorem 5.3 *Let $\{c_x\}_{x\in\mathbb{Z}^2}$ be those of the North-East model. There exists $q_0 \in (0,1)$ and for any $r < \infty$ there exists M_1 such that, for any $M < \min(M_0, M_1)$ and $q \ge q_0$,*

$$\inf_{\Phi \in \mathcal{B}_{r,M}} \operatorname{gap}(\mathcal{L}^\Phi) > 0$$

Remark 5.4. As we will see in the proof of the theorem, the restriction on strength of the interaction comes from two different requirements. The first one is that the finite volume Gibbs measure has the very strong mixing property uniformly in the boundary conditions given in (5.1). That, as we pointed out previously, is guaranteed as long as $M < M_0$. The second one requires that the zeros, which certainly percolate in a robust way w.r.t. the unperturbed measure ν because of the assumption $q \approx 1$, continue to do so even when we switch on the interaction. It is worthwhile to observe that for the one dimensional East model, the first requirement is satisfied for any $M < \infty$ and that the second one is simply not necessary. Therefore for the East model the above theorem should be reformulated as follows.

Theorem 5.5 *Let $\{c_x\}_{x\in\mathbb{Z}}$ be those of the East model. For any finite pair (r, M)*

$$\inf_{\Phi\in\mathcal{B}_{r,M}} \mathrm{gap}(\mathcal{L}^{\Phi}) > 0$$

Proof (of Theorem 5.3). We will follow the pattern of the proof for the non interacting case given in [12] and we will establish the stronger result

$$\sup_{\Lambda\in\mathbb{F}} \gamma(\Lambda) < +\infty, \quad \text{where} \quad \gamma(\Lambda) := \Big(\inf_{\Phi\in\mathcal{B}_{r,M}} \inf_{\tau\in\mathrm{Max}_\Lambda} \mathrm{gap}(\mathcal{L}_\Lambda^{\Phi,\tau})\Big)^{-1} \tag{5.2}$$

provided that $q > q_0$ is large and M is taken sufficiently small. Above Max_Λ denotes the set of configurations in Ω_{Λ^c} which are identically equal to zero on $\partial^*_+\Lambda$. In what follows in order to simplify the notation we will not write the dependence on the boundary condition of the transition rates.

As in [12] the first step consists in proving a certain monotonicity property of $\gamma(\Lambda)$.

Lemma 5.6 *For any $V \subset \Lambda \in \mathbb{F}$,*

$$0 < \gamma(V) \le \gamma(\Lambda) < \infty$$

Proof (Proof of the Lemma). Fix $\Phi \in \mathcal{B}_{r,M}$ and, for any $\xi \in \mathrm{Max}_V$, define the new interaction Φ^ξ as follows:

$$\Phi^\xi_A(\sigma_A) = \begin{cases} 0 & \text{if } A \cap V^c \neq \emptyset \\ \sum_{A' :\, A'\cap V = A} \Phi_{A'}(\sigma_A \cdot \xi_{A'\setminus A}) & \text{if } A \subset V \end{cases}$$

Notice that, by construction,

$$r(\Phi^\xi) \le r(\Phi) \quad \text{and} \quad \sup_x \sum_{A \ni x} \|\Phi^\xi_A\|_\infty \le \|\Phi\|_\infty$$

so that $\Phi^\xi \in \mathcal{B}_{r,M}$. Next observe that the Gibbs measure on Λ with interaction Φ^ξ is simply the product measure

$$\mu^{\Phi^\xi}_\Lambda(\sigma_\Lambda) := \mu^{\Phi,\xi}_V(\sigma_V) \otimes \nu_{\Lambda\setminus V}(\sigma_{\Lambda\setminus V}) \quad \text{on} \quad \Omega_\Lambda = \Omega_V \otimes \Omega_{\Lambda\setminus V}$$

Thus, for any $f \in L^2(\Omega_V, \mu^{\Phi,\xi}_V)$ and $\tau \in \mathrm{Max}_\Lambda$, we can write ($\mathrm{Var}^{\Phi,\tau}_\Lambda \equiv \mathrm{Var}_{\mu^{\Phi,\tau}_\Lambda}$)

$$\mathrm{Var}^{\Phi,\xi}_V(f) = \mathrm{Var}^{\Phi^\xi,\tau}_\Lambda(f) \le \gamma(\Lambda)\, \mathcal{D}^{\Phi^\xi,\tau}_\Lambda(f) \le \gamma(\Lambda)\, \mathcal{D}^{\Phi,\xi}_V(f)$$

where, in the last inequality, we used the fact that, for any $x \in V$ and any $\omega \in \Omega_\Lambda$, $c_{x,\Lambda}(\omega) \le c_{x,V}(\omega)$ because $\xi \in \mathrm{Max}_V$, together with

$$\mathrm{Var}^{\Phi^\xi,\tau}_\Lambda(f \,|\, \{\sigma_y\}_{y\neq x}) = \mathrm{Var}^{\Phi,\xi}_V(f \,|\, \{\sigma_y\}_{y\neq x}).$$

□

Thanks to Lemma 5.6 we need to prove (5.2) only when Λ runs through all possible rectangles. For this purpose our main ingredient will be the bisection technique of [28] which, in its essence, consists in proving a suitable recursion relation between spectral gap on scale $2L$ with that on scale L, combined with the novel idea of considering an accelerated block dynamics which is itself constrained. Such an approach is referred to in [12] as the *Bisection-Constrained* or *B-C approach*.

In order to present it we first need to recall some simple facts from two dimensional percolation.

A *path* is a collection $\{x_0, x_1, \dots, x_n\}$ of distinct points in $\mathbb{Z}^2$ such that $d(x_i, x_{i+1}) = 1$ for all i. A $*$*-path* is a collection $\{x_0, x_1, \dots, x_n\}$ of distinct points in $\mathbb{Z}^2$ such that $x_{i+1} \in \mathcal{N}^*_{x_i}$ for all i. Given a rectangle Λ and a direction $\mathbf{e}_i$, we will say that a path $\{x_0, \dots, x_n\}$ traverses Λ in the i^{th}-direction if $\{x_0, \dots, x_n\} \subset \Lambda$ and x_0, x_n lay on the two opposite sides of Λ orthogonal to $\mathbf{e}_i$.

Definition 5.7. *Given a rectangle Λ and a configuration $\omega \in \Omega_\Lambda$, a path $\{x_0, \dots, x_n\}$ is called a* top-bottom crossing *(*left-right crossing*) if it traverses Λ in the vertical (horizontal) direction and $\omega_{x_i} = 0$ for all $i = 0, \dots, n$. The rightmost (lower-most) such crossings (see [22] page 317) will be denoted by Π_ω.*

Remark 5.8. Given a rectangle Λ and a path Γ traversing Λ in e.g. the vertical direction, let Λ_Γ consists of all the sites in Λ which are in Γ or to the right of it. Then, as remarked in [22], the event $\{\omega : \ \Pi_\omega = \Gamma\}$ depends only on the variables ω_x with $x \in \Lambda_\Gamma$.

We are now ready to start the actual proof of the theorem. At the beginning the method requires a simple geometric result (see [7]) which we now describe.

Let $l_k := (3/2)^{k/2}$, and let $\mathbb{F}_k$ be the set of all rectangles $\Lambda \subset \mathbb{Z}^2$ which, modulo translations and permutations of the coordinates, are contained in $[0, l_{k+1}] \times [0, l_{k+2}]$. The main property of $\mathbb{F}_k$ is that each rectangle in $\mathbb{F}_k \setminus \mathbb{F}_{k-1}$ can be obtained as a "slightly overlapping union" of two rectangles in $\mathbb{F}_{k-1}$.

Lemma 5.9 *For all $k \in \mathbb{Z}_+$, for all $\Lambda \in \mathbb{F}_k \setminus \mathbb{F}_{k-1}$ there exists a finite sequence $\{\Lambda_1^{(i)}, \Lambda_2^{(i)}\}_{i=1}^{s_k}$ in $\mathbb{F}_{k-1}$, where $s_k := \lfloor l_k^{1/3} \rfloor$, such that, letting $\delta_k := \frac{1}{8}\sqrt{l_k} - 2$,*

(i) $\Lambda = \Lambda_1^{(i)} \cup \Lambda_2^{(i)}$,
(ii) $d(\Lambda \setminus \Lambda_1^{(i)}, \Lambda \setminus \Lambda_2^{(i)}) \geq \delta_k$,
(iii) $\left(\Lambda_1^{(i)} \cap \Lambda_2^{(i)}\right) \cap \left(\Lambda_1^{(j)} \cap \Lambda_2^{(j)}\right) = \emptyset$, *if* $i \neq j$.

The *B-C approach* then establishes a simple recursive inequality between the quantity $\gamma_k := \sup_{\Lambda \in \mathbb{F}_k} \gamma(\Lambda)$ on scale k and the same quantity on scale $k-1$ as follows.

Fix $\Lambda \in \mathbb{F}_k \setminus \mathbb{F}_{k-1}$ and write it as $\Lambda = \Lambda_1 \cup \Lambda_2$ with $\Lambda_1, \Lambda_2 \in \mathbb{F}_{k-1}$ satisfying the properties described in Lemma 5.9 above. Without loss of generality we can assume that all the horizontal faces of Λ_1 and of Λ_2 lay on the horizontal faces of Λ except for the face orthogonal to the first direction $\mathbf{e}_1$ and that, along that direction, Λ_1 comes before Λ_2. Set $\Delta \equiv \Lambda_1 \cap \Lambda_2$ and write, for definiteness, $\Delta = [a_1, b_1] \times [a_2, b_2]$. Lemma 5.9 implies that the width of Δ in the first direction, $b_1 - a_1$, is at least δ_k. Set also

$$I \equiv [a_1 + (b_1 - a_1)/2, \ b_1] \times [a_2, b_2]$$

and let $\partial_r I = \{b_1\} \times [a_2, b_2]$ be the right face of I along the first direction.

Definition 5.10. *Given a configuration $\omega \in \Omega$ we will say that ω is I-good iff there exists a top-bottom crossing of I.*

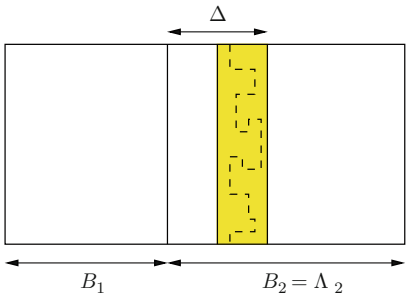


Fig. 1. The rectangle Λ divided into two blocks B_1 and B_2. The grey region is the strip I with a top-bottom crossing.

Given $\tau \in \mathrm{Max}_\Lambda$, we run the following constrained "block dynamics" on Ω_Λ (in what follows, for simplicity, we suppress the index i) with boundary conditions τ and blocks $B_1 := \Lambda_1 \setminus I$, $B_2 := \Lambda_2$. The block B_2 waits a mean one exponential random time and then the current configuration inside it is refreshed with a new one sampled from the Gibbs measure of the block given the previous configuration outside it (and τ outside Λ). The block B_1 does the same but now the configuration is refreshed only if the current configuration ω in B is I-good (see Figure 1).

The generator of the block dynamics applied to f can be written as

$$\mathcal{L}_{\mathrm{block}} f = c_1(\mu^{\Phi,\tau}_{B_1}(f) - f) + \mu^{\Phi,\tau}_{B_2}(f) - f \tag{5.3}$$

and the associated Dirichlet form is

$$\mathcal{D}^{\Phi,\tau}_{\mathrm{block}}(f) = \mu^{\Phi,\tau}_\Lambda \left(c_1 \operatorname{Var}^\Phi_{B_1}(f) + \operatorname{Var}^\Phi_{B_2}(f) \right)$$

where $c_1(\omega)$ is just the indicator of the event that ω is I-good.

Remark 5.11. The reader should keep in mind that the notation

$$\mu^{\Phi,\tau}_\Lambda \left(c_1 \operatorname{Var}^\Phi_{B_1}(f) \right)$$

stands for $\sum_\xi \mu^{\Phi,\tau}_\Lambda(\xi) c_1(\xi) \operatorname{Var}^{\Phi,\xi}_{B_1}(f)$ and that one can imagine the sum restricted to those configurations outside B_1 that coincide with τ outside Λ since otherwise their probability $\mu^{\Phi,\tau}_\Lambda(\xi)$ is zero.

In order to study the mixing property of the chain we need the following two lemmas.

Lemma 5.12 ([28]) *Fix (r, M) with $M < M_0$. Then, for any $\Phi \in \mathcal{B}_{r,M}$,*

$$\begin{aligned} \sup_{\tau'} |\mu_{B_1}^{\Phi,\tau'}(g) - \mu_{\Lambda}^{\Phi,\tau}(g)| &\leq \lambda_k \|g\|_\infty \quad \forall g : \Omega_{B_2^c} \mapsto \mathbb{R} \\ \sup_{\tau'} |\mu_{B_2}^{\Phi,\tau'}(g) - \mu_{\Lambda}^{\Phi,\tau}(g)| &\leq \lambda_k \|g\|_\infty \quad \forall g : \Omega_{B_1^c} \mapsto \mathbb{R} \end{aligned} \tag{5.4}$$

where $\lambda_k := Krl_{k+1}e^{-m\delta_k/2}$ and the constants K, m are given in (5.1).

Lemma 5.13 *There exists $q_0 \in (0,1)$ and for any $r < \infty$ there exists M_1 such that, for any $M < \min(M_0, M_1)$ and $q \geq q_0$,*

$$\varepsilon_k := \max_{\Phi \in \mathcal{B}_{r,M}} \max_{\tau} \mu_{B_2}^{\Phi,\tau}(\omega \text{ is not } I\text{-good }) \leq e^{-\delta_k}.$$

Proof. It follows immediately from standard percolation arguments together with

$$\sup_{\Phi \in \mathcal{B}_{r,M}} \sup_{\tau} \mu_{\{x\}}^{\Phi,\tau}(\sigma_x = 1) \leq (1-q)e^{2M}$$

□

We can now state the main consequence of Lemma 5.12, 5.13.

Proposition 5.14 *There exists $q_0 \in (0,1)$ and for any $r < \infty$ there exists M_1 such that, for any $M < \min(M_0, M_1)$ and $q \geq q_0$,*

$$\gamma_{\text{block}}^{(k)} := \sup_{\Phi \in \mathcal{B}_{r,M}} \sup_{\tau \in \text{Max}_\Lambda} \left(\text{gap}(\mathcal{L}_{\text{block}}^{\Phi,\tau})\right)^{-1} \leq \left(1 - 8\sqrt{2\lambda_k + \varepsilon_k}\right)^{-1} \tag{5.5}$$

for all k so large that the r.h.s. of (5.5) *is smaller than* 2.

Proof (Proof of the proposition). Fix r, M and Φ as prescribed, let $\tau \in \text{Max}_\Lambda$ and, in order to simplify the notation, drop all the superscripts Φ, τ. Let $f : \Omega_\Lambda \mapsto \mathbb{R}$ be a mean zero function, the eigenvalue equation associated to the generator (5.3) is

$$c_1(\mu_{B_1}(f) - f) + \mu_{B_2}(f) - f = \lambda f \tag{5.6}$$

By construction $\lambda \geq -2$.

Assume that $\lambda > -1 + \sqrt{\lambda_k}$ since otherwise there is nothing to be proved. By applying μ_{B_1} to both sides of (5.6) and using (5.4) we obtain

$$(1+\lambda)\mu_{B_1} f = \mu_{B_1}\big(\mu_{B_2}(f)\big) \quad \Rightarrow \quad \|\mu_{B_1}(f)\|_\infty \leq \sqrt{\lambda_k}\, \|\mu_{B_2}(f)\|_\infty \tag{5.7}$$

If we rewrite (5.6) as

$$f = \frac{1}{1+\lambda+c_1}\mu_{B_2}(f) + \frac{c_1}{1+\lambda+c_1}\mu_{B_1}(f)$$

and apply μ_{B_2} to both sides, by using (5.7) together with the assumption $\lambda > -1 + \sqrt{\lambda_k}$, we get

$$\begin{aligned}\|\mu_{B_2}(f)\|_\infty &\le \|\mu_{B_2}(f)\|_\infty \, \|\mu_{B_2}(\frac{1}{1+\lambda+c_1})\|_\infty \\ &\quad +\lambda_k \|\frac{c_1}{1+\lambda+c_1}\|_\infty \|\mu_{B_1}(f)\|_\infty \\ &\le \|\mu_{B_2}(f)\|_\infty \left(\|\mu_{B_2}(\frac{1}{1+\lambda+c_1})\|_\infty + \sqrt{\lambda_k} \right) \end{aligned} \tag{5.8}$$

which is possible only if

$$\|\mu_{B_2}(\frac{1}{1+\lambda+c_1})\|_\infty \ge 1 - \sqrt{\lambda_k}$$

i.e.

$$\lambda \le -1 + 8\sqrt{2\lambda_k + \epsilon_k}$$

and the proof is complete. □

By writing down the standard Poincaré inequality for the block auxiliary chain, we get that for any f

$$\mathrm{Var}_\Lambda^{\Phi,\tau}(f) \le \gamma_{\mathrm{block}}^{(k)} \, \mu_\Lambda^{\Phi,\tau}\Big(c_1 \, \mathrm{Var}_{B_1}^{\Phi}(f) + \mathrm{Var}_{B_2}^{\Phi}(f) \Big) \tag{5.9}$$

The second term in the r.h.s. of (5.9), using the definition of γ_k and the fact that $B_2 = \Lambda_2 \in \mathbb{F}_{k-1}$ is bounded from above by

$$\mu_\Lambda^{\Phi,\tau}\Big(\mathrm{Var}_{B_2}^{\Phi}(f) \Big) \le \gamma_{k-1} \sum_{x \in B_2} \mu_\Lambda^{\Phi,\tau}\big(c_{x,B_2} \, \mathrm{Var}_x^{\Phi}(f) \big) \tag{5.10}$$

Notice that, by construction, for all $x \in B_2$ and all ω, $c_{x,B_2}(\omega) = c_{x,\Lambda}(\omega)$. Therefore the term $\sum_{x\in B_2} \mu_\Lambda^{\Phi,\tau}\big(c_{x,B_2} \, \mathrm{Var}_x^{\Phi}(f)\big)$ is nothing but the contribution carried by the set B_2 to the full Dirichlet form $\mathcal{D}_\Lambda^{\Phi,\tau}(f)$.

Next we examine the more complicate term $\mu_\Lambda^{\Phi,\tau}\Big(c_1 \, \mathrm{Var}_{B_1}^{\Phi}(f) \Big)$. For any ω such that there exists a rightmost crossing Π_ω in I denote by Λ_ω the set of all sites in Λ which are to the *left* of Π_ω. Since $\mathrm{Var}_{B_1}^{\Phi}(f)$ depends only on $\omega_{\Lambda\setminus B_1}$ and, for any top-bottom crossing Γ of I, $\mathbb{1}_{\{\Pi_\omega = \Gamma\}}$ does not depend on the variables ω's to the left of Γ, we can write

$$\mu_{\Lambda}^{\Phi,\tau}\Big(c_1\,\mathrm{Var}_{B_1}^{\Phi}(f)\Big)=\mu_{\Lambda}^{\Phi,\tau}\Big(\mathbb{1}_{\{\exists\,\Pi_\omega\text{ in }I\}}\mu_{\Lambda_\omega}^{\Phi}\big(\mathrm{Var}_{B_1}^{\Phi}(f)\big)\Big)\tag{5.11}$$

The convexity of the variance implies that

$$\mu_{\Lambda_\omega}^{\Phi}\big(\mathrm{Var}_{B_1}^{\Phi}(f)\big)\le\mathrm{Var}_{\Lambda_\omega}^{\Phi}(f)$$

where it is understood that the r.h.s. depends on the variables in Π_ω and to the right of it. The key observation at this stage, which explains the role and the need of the event $\{\exists\ \Pi_\omega \text{ in } I\}$, is the following. For any ω such that Π_ω exists the variance $\mathrm{Var}_{\Lambda_\omega}^{\Phi}(f)$ is computed with boundary conditions (τ outside Λ and $\omega_{\Lambda\setminus\Lambda_\omega}$) which belong to $\mathrm{Max}_{\Lambda_\omega}$. Therefore we can bound it from above using the Poincaré inequality by

$$\mathrm{Var}_{\Lambda_\omega}^{\Phi}(f)\le\gamma(\Lambda_\omega)\mathcal{D}_{\Lambda_\omega}^{\Phi}(f)\le\gamma(B_1\cup I)\mathcal{D}_{\Lambda_\omega}^{\Phi}(f)$$

where we used Lemma 5.6 together with the observation that $\Lambda_\omega\subset B_1\cup I=\Lambda_1$. In conclusion

$$\begin{aligned}&\mu_{\Lambda}^{\Phi,\tau}\Big(\mathbb{1}_{\{\exists\,\Pi_\omega\text{ in }I\}}\mu_{\Lambda_\omega}^{\Phi}\big(\mathrm{Var}_{B_1}^{\Phi}(f)\big)\Big)\\&\le\gamma(\Lambda_1)\mu_{\Lambda}^{\Phi,\tau}\Big(\mathbb{1}_{\{\exists\,\Pi_\omega\text{ in }I\}}\mathcal{D}_{\Lambda_\omega}^{\Phi}(f)\Big)\\&\le\gamma(\Lambda_1)\mu_{\Lambda}^{\Phi,\tau}\Big(\mathbb{1}_{\{\exists\,\Pi_\omega\text{ in }I\}}\sum_{x\in\Lambda_\omega}c_{x,\Lambda_\omega}\,\mathrm{Var}_x^{\Phi}(f)\Big)\\&\le\gamma(\Lambda_1)\mu_{\Lambda}^{\Phi,\tau}\Big(\sum_{x\in\Lambda_1}c_{x,\Lambda}\,\mathrm{Var}_x^{\Phi}(f)\Big)\end{aligned}$$

because, by construction, for every ω such that there exists Π_ω in I

$$c_{x,\Lambda_\omega}(\omega)=c_{x,\Lambda}(\omega)\quad\forall x\in\Lambda_\omega\,.\tag{5.12}$$

If we finally plug (5.1) into the r.h.s. of (5.11) and recall that $\Lambda_1\in\mathcal{F}_{k-1}$, we obtain

$$\mu_{\Lambda}^{\Phi,\tau}\Big(c_1\,\mathrm{Var}_{B_1}^{\Phi}(f)\Big)\le\gamma_{k-1}\,\mu_{\Lambda}^{\Phi,\tau}\big(\sum_{x\in\Lambda_1}c_{x,\Lambda}\,\mathrm{Var}_x^{\Phi}(f)\big)\tag{5.13}$$

In conclusion we have shown that

$$\mathrm{Var}_{\Lambda}^{\Phi,\tau}(f)\le\gamma_{\mathrm{block}}^{(k)}\gamma_{k-1}\Big(\mathcal{D}_{\Lambda}^{\Phi,\tau}(f)+\sum_{x\in\Delta}\mu_{\Lambda}^{\Phi,\tau}\big(c_{x,\Lambda}\,\mathrm{Var}_x(f)\big)\Big)\tag{5.14}$$

Averaging over the $s_k=\lfloor l_k^{1/3}\rfloor$ possible choices of the sets Λ_1,Λ_2 gives

$$\mathrm{Var}_{\Lambda}(f)\le\gamma_{\mathrm{block}}^{(k)}\gamma_{k-1}(1+\frac{1}{s_k})\mathcal{D}_{\Lambda}(f)\tag{5.15}$$

which implies that

$$\gamma_k \le (1+\frac{1}{s_k})\gamma_{\text{block}}^{(k)}\gamma_{k-1} \le \gamma_{k_0} \prod_{j=k_0}^{k}(1+\frac{1}{s_j})\gamma_{\text{block}}^{(j)} \tag{5.16}$$

where k_0 is the smallest integer such that $\gamma_{\text{block}}^{(k_0)} < 2$. If we now recall the expression (5.5) for $\gamma_{\text{block}}^{(j)}$ together with Lemma 5.12 and 5.13, we immediately conclude that the product $\prod_{j=k_0}^{\infty}\gamma_{\text{block}}^{(j)}(1+\frac{1}{s_j})$ is bounded. □

6 One Spin Facilitated Model on a General Graph

In this section we prove our second set of new results by examining the *one spin facilitated* model (FA-1f in short) on a general connected graph $\mathcal{G} = (V, E)$. Our motivation comes from some unpublished speculation by D. Aldous [3] that, in this general setting, the FA-1f may serve as an algorithm for information storage in dynamic graphs.

We begin by discussing the finite setting. Let r be one of the vertices and $\mathcal{T}$ be a rooted spanning tree of $\mathcal{G}$ with root r. On $\Omega = \{0,1\}^V$ consider the FA-1f constraints:

$$\begin{cases} c_{x,\mathcal{G}}(\omega) = 1 & \text{if } \omega_y = 0 \text{ for some neighbor } y \text{ of } x \\ 0 & \text{otherwise} \end{cases} \tag{6.1}$$

and let $\hat{c}_{x,\mathcal{G}} = c_{x,\mathcal{G}}$ if $x \neq r$ and $\hat{c}_{r,\mathcal{G}} \equiv 1$. Let $\hat{\mathcal{L}}$ be the corresponding Markov generator and notice that associated Markov chain is ergodic since the vertex r is unconstrained. For shortness we will refer in the sequel to $\hat{\mathcal{L}}$ as the ($\mathcal{G}$, r, FA-1f) model. Our first result reads as follows.

Theorem 6.1

$$\text{gap}(\mathcal{G}, r, \text{FA-1f}) \ge \text{gap}(\mathbb{Z}, \text{East})$$

Proof. By monotonicity $\hat{c}_{x,\mathcal{G}}(\omega) \ge \hat{c}_{x,\mathcal{T}}(\omega)$ and therefore gap($\mathcal{G}, r$, FA-1f) $\ge$ gap($\mathcal{T}, r$, FA-1f). We can push the monotonicity argument a bit further and consider the following ($\mathcal{T}, r$, East) model:

$$\tilde{c}_{x,\mathcal{T}}(\omega) = \begin{cases} 1 \text{ if either } x = r \text{ or } \omega_y = 0, \text{ where } y \text{ is the ancestor (in } \mathcal{T}) \text{ of } x \\ 0 \text{ otherwise} \end{cases} \tag{6.2}$$

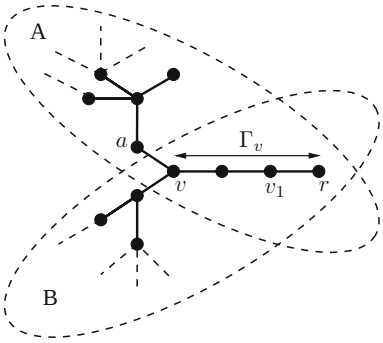


Fig. 2. The subtrees A and B.

Clearly $\hat{c}_{x,\mathcal{T}}(\omega) \geq \tilde{c}_{x,\mathcal{T}}(\omega)$ and therefore $\mathrm{gap}(\mathcal{G}, r, \text{FA-1f}) \geq \mathrm{gap}(\mathcal{T}, r,$ East). We will now proceed to show that

$$\mathrm{gap}(\mathcal{T}, r, \text{East}) \geq \mathrm{gap}(\mathbb{Z}, \text{East}) \tag{6.3}$$

If all the vertices of $\mathcal{T}$ have degree 2 with the exception of the root and the leaves, *i.e.* if $\mathcal{T} \subset \mathbb{Z}$, then (6.3) follows from [12, Lemma 2.11]. Thus let us assume that there exists $x \in \mathcal{T}$ with $\Delta_x \geq 3$ and let us order the vertices of $\mathcal{T}$ by first assigning some arbitrary order to all vertices belonging to any given layer ($\equiv$ same distance from the root) and then declaring $x < y$ iff either $d(x,r) < d(y,r)$ or $d(x,r) = d(y,r)$ and x comes before y in the order assigned to their layer. Let v be equal to the root if $\Delta_r \geq 2$ or equal to the first descendant of r with degree $\Delta_v \geq 3$ otherwise and let $\Gamma_v = \{r, v_1, \ldots, v_k, v\}$ be the path in $\mathcal{T}$ leading from r to v. Let a be a child of v and let $\mathcal{T}_a = (V_a, E_a)$ be the subtree of $\mathcal{T}$ rooted in a. Finally we denote by A and B the two subgraphs of $\mathcal{T}$: $A := \Gamma_v \cup \mathcal{T}_a$, $B := \mathcal{T} \setminus \mathcal{T}_a$. (see Fig. 2).

Lemma 6.2

$$\mathrm{gap}(\mathcal{T}, r, \text{East}) \geq \min\left(\mathrm{gap}(A, r, \text{East}), \mathrm{gap}(B, r, \text{East})\right) \tag{6.4}$$

By recursively applying the above result to A and B separately, we immediately reduce ourselves to the case of a tree $\mathcal{T}' \subset \mathbb{Z}$ and the proof of the theorem is complete. □

Proof (of Lemma 6.2). In $L^2(\Omega, \mu)$ consider the set $\mathcal{H}_B$ of functions f that do not depend on ω_x, $x \in \mathcal{T}_a$. Because of the choice of the constraints $\tilde{c}_{x,\mathcal{T}}(\omega)$, $\mathcal{H}_B$ is an invariant subspace for the generator of the $(\mathcal{T}, r, \text{East})$ model and

$$\inf_{f\in\mathcal{H}_B\mu(f)=0}\frac{\tilde{\mathcal{D}}(f)}{\mathrm{Var}(f)}=\mathrm{gap}(B,r,\mathrm{East}) \tag{6.5}$$

Let us now consider the orthogonal subspace $\mathcal{H}_B^{\perp}$. Any zero mean element $f\in\mathcal{H}_B^{\perp}$ satisfies $\mu_{\mathcal{T}_a}(f)=0$ and therefore we can write

$$\begin{aligned}\mathrm{Var}(f)&=\mu\left(\mathrm{Var}_{\mathcal{T}_a}(f)\right)\le\mu\left(\mathrm{Var}_A(f)\right)\\&\le\mathrm{gap}(A,r,\mathrm{East})^{-1}\sum_{x\in A}\mu\left(\tilde{c}_{x,A}\,\mathrm{Var}_x(f)\right)\\&\le\mathrm{gap}(A,r,\mathrm{East})^{-1}\tilde{\mathcal{D}}(f)\end{aligned}$$

where the first inequality follows from convexity of the variance and the second one is nothing but the Poincaré inequality for the East model in A. The proof of the Lemma follows at once from (6.5). □

Theorem 6.1 has two consequences that will be the content of the following Theorems. The first one deals with the case of an infinite graph. The second one deals with the FA-1f model on general graph $\mathcal{G}$ without the special unblocked vertex r but with the Markov chain restricted to a suitable ergodic component.

Theorem 6.3 *Let $\mathcal{G}_\infty$ be an infinite connected graph of bounded degree and let $\mathcal{L}$ be the generator of the FA-1f model on $\mathcal{G}_\infty$ with constraints $\{c_{x,\mathcal{G}_\infty},\ x\in V_\infty\}$, i.e. no apriori unblocked vertex. Then*

$$\mathrm{gap}(\mathcal{G}_\infty,\mathrm{FA\text{-}1f})\ge\mathrm{gap}(\mathbb{Z},\mathrm{East})$$

Proof. The proof combines Theorem 6.1 together with the finite subgraph approximation described in section 3. □

Theorem 6.4 *Let $\mathcal{G}$ be as in Theorem 6.1 and let $\mathcal{L}^+$ be the FA-1f generator with constraints $\{c_{x,\mathcal{G}}\}_{x\in V}$ on the restricted configuration space $\Omega^+:=\{\eta\in\Omega:\ \sum_{x\in V}(1-\eta_x)\ge 1\}$ equipped with the reversible measure $\mu^+:=\mu(\cdot\,|\,\Omega^+)$. Then*

$$\mathrm{gap}(\mathcal{L}^+)\ge\frac{1}{2^{\mathrm{gap}(\mathbb{Z},\mathrm{East})}}$$

Proof. As in the proof of Theorem 6.1 we can safely assume that $\mathcal{G}$ is a tree $\mathcal{T}$ with root $r\in V$. We extend any $f:\Omega^+\mapsto\mathbb{R}$ to a function $\tilde{f}$ on Ω by setting $\tilde{f}(\eta_y=1\ \forall y)\equiv f(\eta_y=1\ \forall y\neq r,\ \eta_r=0)$. Using Theorem 6.1, we then write

$$\begin{aligned}\mathrm{Var}^+(f)=\mathrm{Var}^+(\tilde{f})&\le\left(\mu(\Omega^+)\right)^{-1}\mathrm{Var}(\tilde{f})\\&\le\left(\mu(\Omega^+)\right)^{-1}\mathrm{gap}(\mathcal{T},r,\mathrm{East})^{-1}\sum_x\mu\left(\hat{c}_{x,\mathcal{T}}\,\mathrm{Var}_x(\tilde{f})\right)\end{aligned}$$

where the constraints $\{\hat{c}_{x,\mathcal{T}}\}_{x\in\mathcal{T}}$ have been defined right after (6.1).

Let us examine a generic term $\mu\left(\hat{c}_{x,\mathcal{T}}\,\mathrm{Var}_x(\tilde{f})\right)$ with $x \neq r$. Remember that $\hat{c}_{x,\mathcal{T}} = c_{x,\mathcal{T}}$ and moreover $c_{x,\mathcal{T}}(\eta) = 0$ if $\eta_y = 1$ for all $y \neq x$. Furthermore, for any η such that there exists $y \neq x$ with $\eta_y = 0$, $\mu^+(\eta_x = 1 \,|\, \{\eta_y\}_{y\neq x}) = p$. In conclusion we have shown that

$$\mu\left(\hat{c}_{x,\mathcal{T}}\,\mathrm{Var}_x(\tilde{f})\right) = \mu(\Omega^+)\mu^+\left(c_{x,\mathcal{T}}\,\mathrm{Var}_x^+(f)\right) \qquad \forall x \neq r \tag{6.6}$$

We now examine the dangerous term $\mu\left(\hat{c}_{r,\mathcal{T}}\,\mathrm{Var}_r(\tilde{f})\right) = \mu\left(\mathrm{Var}_r(\tilde{f})\right)$. Because of the definition of $\tilde{f}$ we can safely rewrite it as

$$\mu\left(\mathrm{Var}_r(\tilde{f})\right) = \mu\left(\chi_{\{\exists\, y\neq r:\, \eta_y=0\}}\,\mathrm{Var}_r(f)\right)$$

Let us order the vertices of the tree $\mathcal{T}$ starting from the *furthermost ones* by first assigning some arbitrary order to all vertices belonging to any given layer ($\equiv$ same distance from the root) and then declaring $x < y$ iff either $d(x,r) > d(y,r)$ or $d(x,r) = d(y,r)$ and x comes before y in the order assigned to their layer. Next, for any η such that $\eta_y = 0$ for some $y \neq r$, define $\xi = \min\{y :\ \eta_y = 0\}$ and let $\mathcal{T}_\xi := \{z \in \mathcal{T} :\ z > \xi\}$ (see Fig. 3).

Notice that the subgraph $\mathcal{T}_\xi$ is again a tree and we define its root to be the ancestor v of ξ in $\mathcal{T}$. Then, using convexity of the variance, we can write

$$\mu\left(\chi_{\{\exists\, y\neq r:\, \eta_y=0\}}\,\mathrm{Var}_r(f)\right) = \mu\left(\chi_{\xi\neq r}\mu\left(\mathrm{Var}_r(f)\,|\,\xi\right)\right) \leq \mu\left(\chi_{\xi\neq r}\,\mathrm{Var}_{\mathcal{T}_\xi}(f)\right)$$

In order to bound from above $\mathrm{Var}_{\mathcal{T}_\xi}(f)$ we apply the Poincaré inequality in $\mathcal{T}_\xi$ with constraints $\{\hat{c}_{z,\mathcal{T}_\xi}\}$ and root v together with Theorem 6.1:

$$\mathrm{Var}_{\mathcal{T}_\xi}(f) \leq \mathrm{gap}(\mathbb{Z}, \mathrm{East})^{-1} \sum_{z\in\mathcal{T}_\xi} \mu_{\mathcal{T}_\xi}\left(\hat{c}_{z,\mathcal{T}_\xi}\,\mathrm{Var}_z(f)\right)$$

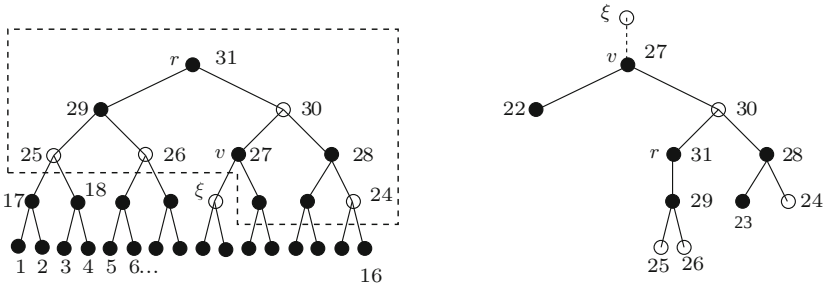


Fig. 3. An example of a tree $\mathcal{T}$ on the left, with a choice of an ordering. The dotted line delimitates the subtree $\mathcal{T}_\xi$ that we reproduce on the right, with root v.

Notice that, by construction, $\hat{c}_{z,\mathcal{T}_\xi}(\eta) = c_{z,\mathcal{T}}(\eta)$ for any $z \in \mathcal{T}_\xi$, including the root v of $\mathcal{T}_\xi$ where $\hat{c}_{v,\mathcal{T}_\xi}(\eta) = 1$ by definition and $c_{v,\mathcal{T}}(\eta) = 1$ because $\eta_\xi = 0$. Putting all together we conclude that

$$\mu\big(\chi_{\{\exists\, y\neq r:\, \eta_y=0\}}\mathrm{Var}_r(f)\big)\leq \mathrm{gap}(\mathbb{Z},\mathrm{East})^{-1}\sum_{x\in\mathcal{T}}\mu\big(\chi_{\{\exists\, y\neq r\eta_y=0\}}c_{x,\mathcal{T}}\mathrm{Var}_x(f)\big)$$
$$\leq \mathrm{gap}(\mathbb{Z},\mathrm{East})^{-1}\mu(\Omega^+)\sum_{x\in\mathcal{T}}\mu^+\left(c_{x,\mathcal{T}}\,\mathrm{Var}_x^+(f)\right)$$

where we have used once more the observation before (6.6) to write

$$c_{x,\mathcal{T}}\,\mathrm{Var}_x(f) = c_{x,\mathcal{T}}\,\mathrm{Var}_x^+(f).$$

If we now combine the previous bounds, we get

$$\mathrm{Var}^+(f) \leq 2\,(\mathrm{gap}(\mathbb{Z},\mathrm{East}))^{-2}\sum_{x\in\mathcal{T}}\mu^+\left(c_{x,\mathcal{T}}\,\mathrm{Var}_x^+(f)\right)$$

and the proof is complete. □

Acknowledgement. F. Martinelli would like to warmly thank Roman Kotecký for the very nice invitation to lecture at the Prague Summer School on *Mathematical Statistical Mechanics*. We also acknowledge P. Sollich for useful discussions and comments on the topics of this paper.

References

1. A. Asselah, P. Dai Pra *Quasi-stationary measures for conservative dynamics in the infinite lattice*, Ann. Probab. **29** (2001), no. 4, 1733–1754.
2. M. Aizenman and J. L. Lebowitz, *Metastability effects in bootstrap percolation*, J. Phys. A **21** (1988), no. 19, 3801–3813. MR MR968311 (90e:82047)
3. D. Aldous, slides available at http://www.math.ucsd.edu/~williams/diaconis/aldous-ising.pdf
4. D. Aldous and P. Diaconis, *The asymmetric one-dimensional constrained Ising model: rigorous results*, J. Statist. Phys. **107** (2002), no. 5-6, 945–975. MR MR1901508 (2003e:82046)
5. L. Berthier and J. P. Garrahan, *Non-topographic description of inherent structure dynamics in glass formers*, J. Chem. Phys. **119** (2003), 4367–4371.
6. L. Berthier, J. P. Garrahan, and S. Whitelam, *Dynamic criticality in glass forming liquids*, Phys. Rev. Lett. (2004), no. 92, 185705–185709.
7. L. Bertini, N. Cancrini, and F. Cesi, *The spectral gap for a Glauber-type dynamics in a continuous gas*, Ann. Inst. H. Poincaré Probab. Statist. **38** (2002), no. 1, 91–108. MR MR1899231 (2003d:82073)

8. G. Biroli, D. S. Fisher, and C. Toninelli, *On the universality of jamming percolation: reply to the comment of jeng and schwarz*, Phys. Rev. Lett. 2007 in press (preprint cond-mat/0612485).
9. G. Biroli and C. Toninelli, *A new class of cellular automata with a discontinuous transition*, preprint.
10. ———, *Jamming percolation and glassy dynamics*, J. Stat. Phys. (2006) in press, preprint cond-mat/0603860, 2006.
11. N. Cancrini, F. Martinelli, C. Roberto, and C. Toninelli, *Kinetically constrained lattice gases*, in preparation.
12. ———, *Kinetically constrained spin models*, Probability Theory and Related Fields in press.
13. ———, *Relaxation times of kinetically constrained spin models with glassy dynamics*, Journal of Statistical Mechanics: Theory and Experiment **L03001** (2007).
14. R. Cerf and E. N. M. Cirillo, *Finite size scaling in three-dimensional bootstrap percolation*, Ann. Probab. **27** (1999), no. 4, 1837–1850. MR MR1742890 (2001b:82047)
15. R. Cerf and F. Manzo, *The threshold regime of finite bootstrap percolation*, Stochastic Process. Appl. **101** (2002), no. 1, 69–82.
16. F. Chung, P. Diaconis, and R. Graham, *Combinatorics for the East model*, Adv. in Appl. Math. **27** (2001), no. 1, 192–206. MR MR1835679 (2002g:82035)
17. C.Toninelli, *Bootstrap and jamming percolation*, Complex Systems, Session LXXXV, Lecture Notes of Houches Summer School, Elsevier, Berlin, July 2006.
18. S. Eisinger and J. Jackle, *A hierarchically constrained kinetic ising model*, Z. Phys. B **84** (1991), 115–124.
19. G. H. Fredrickson and H. C. Andersen, *Kinetic ising model of the glass transition*, Phys. Rev. Lett. **53** (1984), 1244–1247.
20. ———, *Facilitated kinetic ising models and the glass transition*, J. Chem. Phys. **83** (1985), 5822–5831.
21. I. S. Graham, M. Grant, and L. Piché, *Model for dynamics of structural glasses*, Phys. Rev. E **55** (1997), 2132–2144.
22. G. Grimmett, *Percolation*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 321, Springer-Verlag, Berlin, 1999. MR MR1707339 (2001a:60114)
23. P. Harrowell, *Visualizing the collective motion responsible for α and β relaxations in a model glass*, Phys. Rev. E **48** (1993), 4359–4363.
24. A. E. Holroyd, *Sharp metastability threshold for two-dimensional bootstrap percolation*, Probab. Theory Related Fields **125** (2003), no. 2, 195–224. MR MR1961342 (2003k:60257)
25. J. Jackle, F. Mauch, and J. Reiter, *Blocking transitions in lattice spin models with directed kinetic constraints*, Physica A **184** (1992), no. 3–4, 458–476.

26. G. Kordzakhia and S. Lalley, *Ergodicity and mixing properties of the north-east models*, preprint, 2006.
27. T. M. Liggett, *Interacting particle systems*, Springer-Verlag, New York, 1985. MR 86e:60089
28. F. Martinelli, *Lectures on Glauber dynamics for discrete spin models*, Lectures on probability theory and statistics (Saint-Flour, 1997), Springer, Berlin, 1999, pp. 93–191. MR 2002a:60163
29. F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case*, Comm. Math. Phys. **161** (1994), no. 3, 447–486. MR MR1269387 (96c:82040)
30. ———, *Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case*, Comm. Math. Phys. **161** (1994), no. 3, 487–514. MR MR1269388 (96c:82041)
31. F. Martinelli, E. Olivieri, and R. H. Schonmann, *For 2-D lattice spin systems weak mixing implies strong mixing*, Comm. Math. Phys. **165** (1994), no. 1, 33–47. MR MR1298940 (95j:82015)
32. F. Ritort and P. Sollich, *Glassy dynamics of kinetically constrained models*, Advances in Physics **52** (2003), no. 4, 219–342.
33. L. Saloff-Coste, *Lectures on finite Markov chains*, Lectures on probability theory and statistics (Saint-Flour, 1996), Springer, Berlin, 1997, pp. 301–413. MR 99b:60119
34. R. H. Schonmann, *On the behavior of some cellular automata related to bootstrap percolation*, Ann. Probab. **20** (1992), no. 1, 174–193. MR MR1143417 (93b:60231)
35. B. Simon, *The statistical mechanics of lattice gases. Vol. I*, Princeton Series in Physics, Princeton University Press, Princeton, NJ, 1993. MR MR1239893 (95a:82001)
36. P. Sollich and M. R. Evans, *Glassy time-scale divergence and anomalous coarsening in a kinetically constrained spin chain*, Phys. Rev. Lett **83** (1999), 3238–3241.

Index

Lecture Notes in Mathematics

For information about earlier volumes
please contact your bookseller or Springer
LNM Online archive: springerlink.com

Vol. 1786: S. D. Cutkosky, Monomialization of Morphisms from 3-Folds to Surfaces (2002)
Vol. 1787: S. Caenepeel, G. Militaru, S. Zhu, Frobenius and Separable Functors for Generalized Module Categories and Nonlinear Equations (2002)
Vol. 1788: A. Vasil'ev, Moduli of Families of Curves for Conformal and Quasiconformal Mappings (2002)
Vol. 1789: Y. Sommerhäuser, Yetter-Drinfel'd Hopf algebras over groups of prime order (2002)
Vol. 1790: X. Zhan, Matrix Inequalities (2002)
Vol. 1791: M. Knebusch, D. Zhang, Manis Valuations and Prüfer Extensions I: A new Chapter in Commutative Algebra (2002)
Vol. 1792: D. D. Ang, R. Gorenflo, V. K. Le, D. D. Trong, Moment Theory and Some Inverse Problems in Potential Theory and Heat Conduction (2002)
Vol. 1793: J. Cortés Monforte, Geometric, Control and Numerical Aspects of Nonholonomic Systems (2002)
Vol. 1794: N. Pytheas Fogg, Substitution in Dynamics, Arithmetics and Combinatorics. Editors: V. Berthé, S. Ferenczi, C. Mauduit, A. Siegel (2002)
Vol. 1795: H. Li, Filtered-Graded Transfer in Using Noncommutative Gröbner Bases (2002)
Vol. 1796: J M. Melenk, hp-Finite Element Methods for Singular Perturbations (2002)
Vol. 1797: B. Schmidt, Characters and Cyclotomic Fields in Finite Geometry (2002)
Vol. 1798: W.M. Oliva, Geometric Mechanics (2002)
Vol. 1799: H. Pajot, Analytic Capacity, Rectifiability, Menger Curvature and the Cauchy Integral (2002)
Vol. 1800: O. Gabber, L. Ramero, Almost Ring Theory (2003)
Vol. 1801: J. Azéma, M. Émery, M. Ledoux, M. Yor (Eds.), Séminaire de Probabilités XXXVI (2003)
Vol. 1802: V. Capasso, E. Merzbach, B. G. Ivanoff, M. Dozzi, R. Dalang, T. Mountford, Topics in Spatial Stochastic Processes. Martina Franca, Italy 2001. Editor: E. Merzbach (2003)
Vol. 1803: G. Dolzmann, Variational Methods for Crystalline Microstructure – Analysis and Computation (2003)
Vol. 1804: I. Cherednik, Ya. Markov, R. Howe, G. Lusztig, Iwahori-Hecke Algebras and their Representation Theory. Martina Franca, Italy 1999. Editors: V. Baldoni, D. Barbasch (2003)
Vol. 1805: F. Cao, Geometric Curve Evolution and Image Processing (2003)
Vol. 1806: H. Broer, I. Hoveijn. G. Lunther, G. Vegter, Bifurcations in Hamiltonian Systems. Computing Singularities by Gröbner Bases (2003)
Vol. 1807: V. D. Milman, G. Schechtman (Eds.), Geometric Aspects of Functional Analysis. Israel Seminar 2000-2002 (2003)
Vol. 1808: W. Schindler, Measures with Symmetry Properties (2003)
Vol. 1809: O. Steinbach, Stability Estimates for Hybrid Coupled Domain Decomposition Methods (2003)
Vol. 1810: J. Wengenroth, Derived Functors in Functional Analysis (2003)
Vol. 1811: J. Stevens, Deformations of Singularities (2003)
Vol. 1812: L. Ambrosio, K. Deckelnick, G. Dziuk, M. Mimura, V. A. Solonnikov, H. M. Soner, Mathematical Aspects of Evolving Interfaces. Madeira, Funchal, Portugal 2000. Editors: P. Colli, J. F. Rodrigues (2003)
Vol. 1813: L. Ambrosio, L. A. Caffarelli, Y. Brenier, G. Buttazzo, C. Villani, Optimal Transportation and its Applications. Martina Franca, Italy 2001. Editors: L. A. Caffarelli, S. Salsa (2003)
Vol. 1814: P. Bank, F. Baudoin, H. Föllmer, L.C.G. Rogers, M. Soner, N. Touzi, Paris-Princeton Lectures on Mathematical Finance 2002 (2003)
Vol. 1815: A. M. Vershik (Ed.), Asymptotic Combinatorics with Applications to Mathematical Physics. St. Petersburg, Russia 2001 (2003)
Vol. 1816: S. Albeverio, W. Schachermayer, M. Talagrand, Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXX-2000. Editor: P. Bernard (2003)
Vol. 1817: E. Koelink, W. Van Assche (Eds.), Orthogonal Polynomials and Special Functions. Leuven 2002 (2003)
Vol. 1818: M. Bildhauer, Convex Variational Problems with Linear, nearly Linear and/or Anisotropic Growth Conditions (2003)
Vol. 1819: D. Masser, Yu. V. Nesterenko, H. P. Schlickewei, W. M. Schmidt, M. Waldschmidt, Diophantine Approximation. Cetraro, Italy 2000. Editors: F. Amoroso, U. Zannier (2003)
Vol. 1820: F. Hiai, H. Kosaki, Means of Hilbert Space Operators (2003)
Vol. 1821: S. Teufel, Adiabatic Perturbation Theory in Quantum Dynamics (2003)
Vol. 1822: S.-N. Chow, R. Conti, R. Johnson, J. Mallet-Paret, R. Nussbaum, Dynamical Systems. Cetraro, Italy 2000. Editors: J. W. Macki, P. Zecca (2003)
Vol. 1823: A. M. Anile, W. Allegretto, C. Ringhofer, Mathematical Problems in Semiconductor Physics. Cetraro, Italy 1998. Editor: A. M. Anile (2003)
Vol. 1824: J. A. Navarro González, J. B. Sancho de Salas, C^∞ – Differentiable Spaces (2003)
Vol. 1825: J. H. Bramble, A. Cohen, W. Dahmen, Multiscale Problems and Methods in Numerical Simulations, Martina Franca, Italy 2001. Editor: C. Canuto (2003)
Vol. 1826: K. Dohmen, Improved Bonferroni Inequalities via Abstract Tubes. Inequalities and Identities of Inclusion-Exclusion Type. VIII, 113 p, 2003.
Vol. 1827: K. M. Pilgrim, Combinations of Complex Dynamical Systems. IX, 118 p, 2003.
Vol. 1828: D. J. Green, Gröbner Bases and the Computation of Group Cohomology. XII, 138 p, 2003.

Vol. 1829: E. Altman, B. Gaujal, A. Hordijk, Discrete-Event Control of Stochastic Networks: Multimodularity and Regularity. XIV, 313 p, 2003.
Vol. 1830: M. I. Gil', Operator Functions and Localization of Spectra. XIV, 256 p, 2003.
Vol. 1831: A. Connes, J. Cuntz, E. Guentner, N. Higson, J. E. Kaminker, Noncommutative Geometry, Martina Franca, Italy 2002. Editors: S. Doplicher, L. Longo (2004)
Vol. 1832: J. Azéma, M. Émery, M. Ledoux, M. Yor (Eds.), Séminaire de Probabilités XXXVII (2003)
Vol. 1833: D.-Q. Jiang, M. Qian, M.-P. Qian, Mathematical Theory of Nonequilibrium Steady States. On the Frontier of Probability and Dynamical Systems. IX, 280 p, 2004.
Vol. 1834: Yo. Yomdin, G. Comte, Tame Geometry with Application in Smooth Analysis. VIII, 186 p, 2004.
Vol. 1835: O.T. Izhboldin, B. Kahn, N.A. Karpenko, A. Vishik, Geometric Methods in the Algebraic Theory of Quadratic Forms. Summer School, Lens, 2000. Editor: J.-P. Tignol (2004)
Vol. 1836: C. Năstăsescu, F. Van Oystaeyen, Methods of Graded Rings. XIII, 304 p, 2004.
Vol. 1837: S. Tavaré, O. Zeitouni, Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXXI-2001. Editor: J. Picard (2004)
Vol. 1838: A.J. Ganesh, N.W. O'Connell, D.J. Wischik, Big Queues. XII, 254 p, 2004.
Vol. 1839: R. Gohm, Noncommutative Stationary Processes. VIII, 170 p, 2004.
Vol. 1840: B. Tsirelson, W. Werner, Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXXII-2002. Editor: J. Picard (2004)
Vol. 1841: W. Reichel, Uniqueness Theorems for Variational Problems by the Method of Transformation Groups (2004)
Vol. 1842: T. Johnsen, A. L. Knutsen, K_3 Projective Models in Scrolls (2004)
Vol. 1843: B. Jefferies, Spectral Properties of Noncommuting Operators (2004)
Vol. 1844: K.F. Siburg, The Principle of Least Action in Geometry and Dynamics (2004)
Vol. 1845: Min Ho Lee, Mixed Automorphic Forms, Torus Bundles, and Jacobi Forms (2004)
Vol. 1846: H. Ammari, H. Kang, Reconstruction of Small Inhomogeneities from Boundary Measurements (2004)
Vol. 1847: T.R. Bielecki, T. Björk, M. Jeanblanc, M. Rutkowski, J.A. Scheinkman, W. Xiong, Paris-Princeton Lectures on Mathematical Finance 2003 (2004)
Vol. 1848: M. Abate, J. E. Fornaess, X. Huang, J. P. Rosay, A. Tumanov, Real Methods in Complex and CR Geometry, Martina Franca, Italy 2002. Editors: D. Zaitsev, G. Zampieri (2004)
Vol. 1849: Martin L. Brown, Heegner Modules and Elliptic Curves (2004)
Vol. 1850: V. D. Milman, G. Schechtman (Eds.), Geometric Aspects of Functional Analysis. Israel Seminar 2002-2003 (2004)
Vol. 1851: O. Catoni, Statistical Learning Theory and Stochastic Optimization (2004)
Vol. 1852: A.S. Kechris, B.D. Miller, Topics in Orbit Equivalence (2004)
Vol. 1853: Ch. Favre, M. Jonsson, The Valuative Tree (2004)
Vol. 1854: O. Saeki, Topology of Singular Fibers of Differential Maps (2004)
Vol. 1855: G. Da Prato, P.C. Kunstmann, I. Lasiecka, A. Lunardi, R. Schnaubelt, L. Weis, Functional Analytic Methods for Evolution Equations. Editors: M. Iannelli, R. Nagel, S. Piazzera (2004)
Vol. 1856: K. Back, T.R. Bielecki, C. Hipp, S. Peng, W. Schachermayer, Stochastic Methods in Finance, Bressanone/Brixen, Italy, 2003. Editors: M. Fritelli, W. Runggaldier (2004)
Vol. 1857: M. Émery, M. Ledoux, M. Yor (Eds.), Séminaire de Probabilités XXXVIII (2005)
Vol. 1858: A.S. Cherny, H.-J. Engelbert, Singular Stochastic Differential Equations (2005)
Vol. 1859: E. Letellier, Fourier Transforms of Invariant Functions on Finite Reductive Lie Algebras (2005)
Vol. 1860: A. Borisyuk, G.B. Ermentrout, A. Friedman, D. Terman, Tutorials in Mathematical Biosciences I. Mathematical Neurosciences (2005)
Vol. 1861: G. Benettin, J. Henrard, S. Kuksin, Hamiltonian Dynamics – Theory and Applications, Cetraro, Italy, 1999. Editor: A. Giorgilli (2005)
Vol. 1862: B. Helffer, F. Nier, Hypoelliptic Estimates and Spectral Theory for Fokker-Planck Operators and Witten Laplacians (2005)
Vol. 1863: H. Führ, Abstract Harmonic Analysis of Continuous Wavelet Transforms (2005)
Vol. 1864: K. Efstathiou, Metamorphoses of Hamiltonian Systems with Symmetries (2005)
Vol. 1865: D. Applebaum, B.V. R. Bhat, J. Kustermans, J. M. Lindsay, Quantum Independent Increment Processes I. From Classical Probability to Quantum Stochastic Calculus. Editors: M. Schürmann, U. Franz (2005)
Vol. 1866: O.E. Barndorff-Nielsen, U. Franz, R. Gohm, B. Kümmerer, S. Thorbjønsen, Quantum Independent Increment Processes II. Structure of Quantum Lévy Processes, Classical Probability, and Physics. Editors: M. Schürmann, U. Franz, (2005)
Vol. 1867: J. Sneyd (Ed.), Tutorials in Mathematical Biosciences II. Mathematical Modeling of Calcium Dynamics and Signal Transduction. (2005)
Vol. 1868: J. Jorgenson, S. Lang, $\mathrm{Pos}_n(\mathrm{R})$ and Eisenstein Series. (2005)
Vol. 1869: A. Dembo, T. Funaki, Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXXIII-2003. Editor: J. Picard (2005)
Vol. 1870: V.I. Gurariy, W. Lusky, Geometry of Müntz Spaces and Related Questions. (2005)
Vol. 1871: P. Constantin, G. Gallavotti, A.V. Kazhikhov, Y. Meyer, S. Ukai, Mathematical Foundation of Turbulent Viscous Flows, Martina Franca, Italy, 2003. Editors: M. Cannone, T. Miyakawa (2006)
Vol. 1872: A. Friedman (Ed.), Tutorials in Mathematical Biosciences III. Cell Cycle, Proliferation, and Cancer (2006)
Vol. 1873: R. Mansuy, M. Yor, Random Times and Enlargements of Filtrations in a Brownian Setting (2006)
Vol. 1874: M. Yor, M. Émery (Eds.), In Memoriam Paul-André Meyer - Séminaire de Probabilités XXXIX (2006)
Vol. 1875: J. Pitman, Combinatorial Stochastic Processes. Ecole d'Eté de Probabilités de Saint-Flour XXXII-2002. Editor: J. Picard (2006)
Vol. 1876: H. Herrlich, Axiom of Choice (2006)
Vol. 1877: J. Steuding, Value Distributions of *L*-Functions (2007)
Vol. 1878: R. Cerf, The Wulff Crystal in Ising and Percolation Models, Ecole d'Eté de Probabilités de Saint-Flour XXXIV-2004. Editor: Jean Picard (2006)

Vol. 1879: G. Slade, The Lace Expansion and its Applications, Ecole d'Eté de Probabilités de Saint-Flour XXXIV-2004. Editor: Jean Picard (2006)
Vol. 1880: S. Attal, A. Joye, C.-A. Pillet, Open Quantum Systems I, The Hamiltonian Approach (2006)
Vol. 1881: S. Attal, A. Joye, C.-A. Pillet, Open Quantum Systems II, The Markovian Approach (2006)
Vol. 1882: S. Attal, A. Joye, C.-A. Pillet, Open Quantum Systems III, Recent Developments (2006)
Vol. 1883: W. Van Assche, F. Marcellàn (Eds.), Orthogonal Polynomials and Special Functions, Computation and Application (2006)
Vol. 1884: N. Hayashi, E.I. Kaikina, P.I. Naumkin, I.A. Shishmarev, Asymptotics for Dissipative Nonlinear Equations (2006)
Vol. 1885: A. Telcs, The Art of Random Walks (2006)
Vol. 1886: S. Takamura, Splitting Deformations of Degenerations of Complex Curves (2006)
Vol. 1887: K. Habermann, L. Habermann, Introduction to Symplectic Dirac Operators (2006)
Vol. 1888: J. van der Hoeven, Transseries and Real Differential Algebra (2006)
Vol. 1889: G. Osipenko, Dynamical Systems, Graphs, and Algorithms (2006)
Vol. 1890: M. Bunge, J. Funk, Singular Coverings of Toposes (2006)
Vol. 1891: J.B. Friedlander, D.R. Heath-Brown, H. Iwaniec, J. Kaczorowski, Analytic Number Theory, Cetraro, Italy, 2002. Editors: A. Perelli, C. Viola (2006)
Vol. 1892: A. Baddeley, I. Bárány, R. Schneider, W. Weil, Stochastic Geometry, Martina Franca, Italy, 2004. Editor: W. Weil (2007)
Vol. 1893: H. Hanßmann, Local and Semi-Local Bifurcations in Hamiltonian Dynamical Systems, Results and Examples (2007)
Vol. 1894: C.W. Groetsch, Stable Approximate Evaluation of Unbounded Operators (2007)
Vol. 1895: L. Molnár, Selected Preserver Problems on Algebraic Structures of Linear Operators and on Function Spaces (2007)
Vol. 1896: P. Massart, Concentration Inequalities and Model Selection, Ecole d'Été de Probabilités de Saint-Flour XXXIII-2003. Editor: J. Picard (2007)
Vol. 1897: R. Doney, Fluctuation Theory for Lévy Processes, Ecole d'Été de Probabilités de Saint-Flour XXXV-2005. Editor: J. Picard (2007)
Vol. 1898: H.R. Beyer, Beyond Partial Differential Equations, On linear and Quasi-Linear Abstract Hyperbolic Evolution Equations (2007)
Vol. 1899: Séminaire de Probabilités XL. Editors: C. Donati-Martin, M. Émery, A. Rouault, C. Stricker (2007)
Vol. 1900: E. Bolthausen, A. Bovier (Eds.), Spin Glasses (2007)
Vol. 1901: O. Wittenberg, Intersections de deux quadriques et pinceaux de courbes de genre 1, Intersections of Two Quadrics and Pencils of Curves of Genus 1 (2007)
Vol. 1902: A. Isaev, Lectures on the Automorphism Groups of Kobayashi-Hyperbolic Manifolds (2007)
Vol. 1903: G. Kresin, V. Maz'ya, Sharp Real-Part Theorems (2007)
Vol. 1904: P. Giesl, Construction of Global Lyapunov Functions Using Radial Basis Functions (2007)
Vol. 1905: C. Prévôt, M. Röckner, A Concise Course on Stochastic Partial Differential Equations (2007)
Vol. 1906: T. Schuster, The Method of Approximate Inverse: Theory and Applications (2007)
Vol. 1907: M. Rasmussen, Attractivity and Bifurcation for Nonautonomous Dynamical Systems (2007)
Vol. 1908: T.J. Lyons, M. Caruana, T. Lévy, Differential Equations Driven by Rough Paths, Ecole d'Été de Probabilités de Saint-Flour XXXIV-2004 (2007)
Vol. 1909: H. Akiyoshi, M. Sakuma, M. Wada, Y. Yamashita, Punctured Torus Groups and 2-Bridge Knot Groups (I) (2007)
Vol. 1910: V.D. Milman, G. Schechtman (Eds.), Geometric Aspects of Functional Analysis. Israel Seminar 2004-2005 (2007)
Vol. 1911: A. Bressan, D. Serre, M. Williams, K. Zumbrun, Hyperbolic Systems of Balance Laws. Cetraro, Italy 2003. Editor: P. Marcati (2007)
Vol. 1912: V. Berinde, Iterative Approximation of Fixed Points (2007)
Vol. 1913: J.E. Marsden, G. Misiołek, J.-P. Ortega, M. Perlmutter, T.S. Ratiu, Hamiltonian Reduction by Stages (2007)
Vol. 1914: G. Kutyniok, Affine Density in Wavelet Analysis (2007)
Vol. 1915: T. Bıyıkoğlu, J. Leydold, P.F. Stadler, Laplacian Eigenvectors of Graphs. Perron-Frobenius and Faber-Krahn Type Theorems (2007)
Vol. 1916: C. Villani, F. Rezakhanlou, Entropy Methods for the Boltzmann Equation. Editors: F. Golse, S. Olla (2008)
Vol. 1917: I. Veselić, Existence and Regularity Properties of the Integrated Density of States of Random Schrödinger (2008)
Vol. 1918: B. Roberts, R. Schmidt, Local Newforms for GSp(4) (2007)
Vol. 1919: R.A. Carmona, I. Ekeland, A. Kohatsu-Higa, J.-M. Lasry, P.-L. Lions, H. Pham, E. Taflin, Paris-Princeton Lectures on Mathematical Finance 2004. Editors: R.A. Carmona, E. Çinlar, I. Ekeland, E. Jouini, J.A. Scheinkman, N. Touzi (2007)
Vol. 1920: S.N. Evans, Probability and Real Trees. Ecole d'Été de Probabilités de Saint-Flour XXXV-2005 (2008)
Vol. 1921: J.P. Tian, Evolution Algebras and their Applications (2008)
Vol. 1922: A. Friedman (Ed.), Tutorials in Mathematical BioSciences IV. Evolution and Ecology (2008)
Vol. 1923: J.P.N. Bishwal, Parameter Estimation in Stochastic Differential Equations (2008)
Vol. 1924: M. Wilson, Littlewood-Paley Theory and Exponential-Square Integrability (2008)
Vol. 1925: M. du Sautoy, L. Woodward, Zeta Functions of Groups and Rings (2008)
Vol. 1926: L. Barreira, V. Claudia, Stability of Nonautonomous Differential Equations (2008)
Vol. 1927: L. Ambrosio, L. Caffarelli, M.G. Crandall, L.C. Evans, N. Fusco, Calculus of Variations and Non-Linear Partial Differential Equations. Cetraro, Italy 2005. Editors: B. Dacorogna, P. Marcellini (2008)
Vol. 1928: J. Jonsson, Simplicial Complexes of Graphs (2008)
Vol. 1929: Y. Mishura, Stochastic Calculus for Fractional Brownian Motion and Related Processes (2008)
Vol. 1930: J M. Urbano, The Method of Intrinsic Scaling. A Systematic Approach to Regularity for Degenerate and Singular PDEs (2008)
Vol. 1931: M. Cowling, E. Frenkel, M. Kashiwara, A. Valette, D.A. Vogan, Jr., N.R. Wallach, Representation

Theory and Complex Analysis. Venice, Italy 2004. Editors: E.C. Tarabusi, A. D'Agnolo, M. Picardello (2008)
Vol. 1932: A.A. Agrachev, A.S. Morse, E.D. Sontag, H.J. Sussmann, V.I. Utkin, Nonlinear and Optimal Control Theory. Cetraro, Italy 2004. Editors: P. Nistri, G. Stefani (2008)
Vol. 1933: M. Petkovic, Point Estimation of Root Finding Methods (2008)
Vol. 1934: C. Donati-Martin, M. Émery, A. Rouault, C. Stricker (Eds.), Séminaire de Probabilités XLI (2008)
Vol. 1935: A. Unterberger, Alternative Pseudodifferential Analysis (2008)
Vol. 1936: P. Magal, S. Ruan (Eds.), Structured Population Models in Biology and Epidemiology (2008)
Vol. 1937: G. Capriz, P. Giovine, P.M. Mariano (Eds.), Mathematical Models of Granular Matter (2008)
Vol. 1938: D. Auroux, F. Catanese, M. Manetti, P. Seidel, B. Siebert, I. Smith, G. Tian, Symplectic 4-Manifolds and Algebraic Surfaces. Cetraro, Italy 2003. Editors: F. Catanese, G. Tian (2008)
Vol. 1939: D. Boffi, F. Brezzi, L. Demkowicz, R.G. Durán, R.S. Falk, M. Fortin, Mixed Finite Elements, Compatibility Conditions, and Applications. Cetraro, Italy 2006. Editors: D. Boffi, L. Gastaldi (2008)
Vol. 1940: J. Banasiak, V. Capasso, M.A.J. Chaplain, M. Lachowicz, J. Miękisz, Multiscale Problems in the Life Sciences. From Microscopic to Macroscopic. Będlewo, Poland 2006. Editors: V. Capasso, M. Lachowicz (2008)
Vol. 1941: S.M.J. Haran, Arithmetical Investigations. Representation Theory, Orthogonal Polynomials, and Quantum Interpolations (2008)
Vol. 1942: S. Albeverio, F. Flandoli, Y.G. Sinai, SPDE in Hydrodynamic. Recent Progress and Prospects. Cetraro, Italy 2005. Editors: G. Da Prato, M. Röckner (2008)
Vol. 1943: L.L. Bonilla (Ed.), Inverse Problems and Imaging. Martina Franca, Italy 2002 (2008)
Vol. 1944: A. Di Bartolo, G. Falcone, P. Plaumann, K. Strambach, Algebraic Groups and Lie Groups with Few Factors (2008)
Vol. 1945: F. Brauer, P. van den Driessche, J. Wu (Eds.), Mathematical Epidemiology (2008)
Vol. 1946: G. Allaire, A. Arnold, P. Degond, T.Y. Hou, Quantum Transport. Modelling, Analysis and Asymptotics. Cetraro, Italy 2006. Editors: N.B. Abdallah, G. Frosali (2008)
Vol. 1947: D. Abramovich, M. Mariño, M. Thaddeus, R. Vakil, Enumerative Invariants in Algebraic Geometry and String Theory. Cetraro, Italy 2005. Editors: K. Behrend, M. Manetti (2008)
Vol. 1948: F. Cao, J-L. Lisani, J-M. Morel, P. Musé, F. Sur, A Theory of Shape Identification (2008)
Vol. 1949: H.G. Feichtinger, B. Helffer, M.P. Lamoureux, N. Lerner, J. Toft, Pseudo-Differential Operators. Quantization and Signals. Cetraro, Italy 2006. Editors: L. Rodino, M.W. Wong (2008)
Vol. 1950: M. Bramson, Stability of Queueing Networks, Ecole d'Eté de Probabilités de Saint-Flour XXXVI-2006 (2008)
Vol. 1951: A. Moltó, J. Orihuela, S. Troyanski, M. Valdivia, A Non Linear Transfer Technique for Renorming (2009)
Vol. 1952: R. Mikhailov, I.B.S. Passi, Lower Central and Dimension Series of Groups (2009)
Vol. 1953: K. Arwini, C.T.J. Dodson, Information Geometry (2008)
Vol. 1954: P. Biane, L. Bouten, F. Cipriani, N. Konno, N. Privault, Q. Xu, Quantum Potential Theory. Editors: U. Franz, M. Schuermann (2008)
Vol. 1955: M. Bernot, V. Caselles, J.-M. Morel, Optimal Transportation Networks (2009)
Vol. 1956: C.H. Chu, Matrix Convolution Operators on Groups (2008)
Vol. 1957: A. Guionnet, On Random Matrices: Macroscopic Asymptotics, Ecole d'Eté de Probabilités de Saint-Flour XXXVI-2006 (2009)
Vol. 1958: M.C. Olsson, Compactifying Moduli Spaces for Abelian Varieties (2008)
Vol. 1959: Y. Nakkajima, A. Shiho, Weight Filtrations on Log Crystalline Cohomologies of Families of Open Smooth Varieties (2008)
Vol. 1960: J. Lipman, M. Hashimoto, Foundations of Grothendieck Duality for Diagrams of Schemes (2009)
Vol. 1961: G. Buttazzo, A. Pratelli, S. Solimini, E. Stepanov, Optimal Urban Networks via Mass Transportation (2009)
Vol. 1962: R. Dalang, D. Khoshnevisan, C. Mueller, D. Nualart, Y. Xiao, A Minicourse on Stochastic Partial Differential Equations (2009)
Vol. 1963: W. Siegert, Local Lyapunov Exponents (2009)
Vol. 1964: W. Roth, Operator-Valued Measures and Integrals for Cone-Valued Functions (2009)
Vol. 1965: C. Chidume, Geometric Properties of Banach Spaces and Nonlinear Iterations (2009)
Vol. 1966: D. Deng, Y. Han, Harmonic Analysis on Spaces of Homogeneous Type (2009)
Vol. 1967: Benoit Fresse, Modules over Operads and Functors (2009)
Vol. 1968: R. Weissauer, Endoscopy for GSP(4) and the Cohomology of Siegel Modular Threefolds (2009)
Vol. 1969: B. Roynette, M. Yor, Penalising Brownian Paths (2009)
Vol. 1970: M. Biskup, A. Bovier, F. den Hollander, D. Ioffe, F. Martinelli, K. Netočný, F. Toninelli, Methods of Contemporary Mathematical Statistical Physics. Editor: R. Kotecký (2009)
Vol. 1971: L. Saint-Raymond, Hydrodynamic Limits of the Boltzmann Equation (2009)
Vol. 1972: T. Mochizuki, Donaldson Type Invariants for Algebraic Surfaces (2009)

Recent Reprints and New Editions

Vol. 1702: J. Ma, J. Yong, Forward-Backward Stochastic Differential Equations and their Applications. 1999 – Corr. 3rd printing (2007)
Vol. 830: J.A. Green, Polynomial Representations of GL_n, with an Appendix on Schensted Correspondence and Littelmann Paths by K. Erdmann, J.A. Green and M. Schoker 1980 – 2nd corr. and augmented edition (2007)
Vol. 1693: S. Simons, From Hahn-Banach to Monotonicity (Minimax and Monotonicity 1998) – 2nd exp. edition (2008)
Vol. 470: R.E. Bowen, Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. With a preface by D. Ruelle. Edited by J.-R. Chazottes. 1975 – 2nd rev. edition (2008)
Vol. 523: S.A. Albeverio, R.J. Høegh-Krohn, S. Mazzucchi, Mathematical Theory of Feynman Path Integral. 1976 – 2nd corr. and enlarged edition (2008)
Vol. 1764: A. Cannas da Silva, Lectures on Symplectic Geometry 2001 – Corr. 2nd printing (2008)

LECTURE NOTES IN MATHEMATICS

Springer

Edited by J.-M. Morel, F. Takens, B. Teissier, P.K. Maini

Editorial Policy (for Multi-Author Publications: Summer Schools/Intensive Courses)

1. Lecture Notes aim to report new developments in all areas of mathematics and their applications - quickly, informally and at a high level. Mathematical texts analysing new developments in modelling and numerical simulation are welcome. Manuscripts should be reasonably self-contained and rounded off. Thus they may, and often will, present not only results of the author but also related work by other people. They should provide sufficient motivation, examples and applications. There should also be an introduction making the text comprehensible to a wider audience. This clearly distinguishes Lecture Notes from journal articles or technical reports which normally are very concise. Articles intended for a journal but too long to be accepted by most journals, usually do not have this "lecture notes" character.
2. In general SUMMER SCHOOLS and other similar INTENSIVE COURSES are held to present mathematical topics that are close to the frontiers of recent research to an audience at the beginning or intermediate graduate level, who may want to continue with this area of work, for a thesis or later. This makes demands on the didactic aspects of the presentation. Because the subjects of such schools are advanced, there often exists no textbook, and so ideally, the publication resulting from such a school could be a first approximation to such a textbook. Usually several authors are involved in the writing, so it is not always simple to obtain a unified approach to the presentation.

   For prospective publication in LNM, the resulting manuscript should not be just a collection of course notes, each of which has been developed by an individual author with little or no co-ordination with the others, and with little or no common concept. The subject matter should dictate the structure of the book, and the authorship of each part or chapter should take secondary importance. Of course the choice of authors is crucial to the quality of the material at the school and in the book, and the intention here is not to belittle their impact, but simply to say that the book should be planned to be written by these authors jointly, and not just assembled as a result of what these authors happen to submit.

   This represents considerable preparatory work (as it is imperative to ensure that the authors know these criteria before they invest work on a manuscript), and also considerable editing work afterwards, to get the book into final shape. Still it is the form that holds the most promise of a successful book that will be used by its intended audience, rather than yet another volume of proceedings for the library shelf.
3. Manuscripts should be submitted either online at www.editorialmanager.com/lnm/ to Springer's mathematics editorial, or to one of the series editors. Volume editors are expected to arrange for the refereeing, to the usual scientific standards, of the individual contributions. If the resulting reports can be forwarded to us (series editors or Springer) this is very helpful. If no reports are forwarded or if other questions remain unclear in respect of homogeneity etc, the series editors may wish to consult external referees for an overall evaluation of the volume. A final decision to publish can be made only on the basis of the complete manuscript; however a preliminary decision can be based on a pre-final or incomplete manuscript. The strict minimum amount of material that will be considered should include a detailed outline describing the planned contents of each chapter.

   Volume editors and authors should be aware that incomplete or insufficiently close to final manuscripts almost always result in longer evaluation times. They should also be aware that parallel submission of their manuscript to another publisher while under consideration for LNM will in general lead to immediate rejection.

4. Manuscripts should in general be submitted in English. Final manuscripts should contain at least 100 pages of mathematical text and should always include
   - a general table of contents;
   - an informative introduction, with adequate motivation and perhaps some historical remarks: it should be accessible to a reader not intimately familiar with the topic treated;
   - a global subject index: as a rule this is genuinely helpful for the reader.

   Lecture Notes volumes are, as a rule, printed digitally from the authors' files. We strongly recommend that all contributions in a volume be written in the same LaTeX version, preferably LaTeX2e. To ensure best results, authors are asked to use the LaTeX2e style files available from Springer's web-server at

   ftp://ftp.springer.de/pub/tex/latex/svmonot1/ (for monographs) and
   ftp://ftp.springer.de/pub/tex/latex/svmultt1/ (for summer schools/tutorials).

   Additional technical instructions are available on request from: lnm@springer.com.
5. Careful preparation of the manuscripts will help keep production time short besides ensuring satisfactory appearance of the finished book in print and online. After acceptance of the manuscript authors will be asked to prepare the final LaTeX source files and also the corresponding dvi-, pdf- or zipped ps-file. The LaTeX source files are essential for producing the full-text online version of the book. For the existing online volumes of LNM see: http://www.springerlink.com/openurl.asp?genre=journal&issn=0075-8434.

   The actual production of a Lecture Notes volume takes approximately 12 weeks.
6. Volume editors receive a total of 50 free copies of their volume to be shared with the authors, but no royalties. They and the authors are entitled to a discount of 33.3% on the price of Springer books purchased for their personal use, if ordering directly from Springer.
7. Commitment to publish is made by letter of intent rather than by signing a formal contract. Springer-Verlag secures the copyright for each volume. Authors are free to reuse material contained in their LNM volumes in later publications: a brief written (or e-mail) request for formal permission is sufficient.

Addresses:

Professor J.-M. Morel, CMLA,
École Normale Supérieure de Cachan,
61 Avenue du Président Wilson,
94235 Cachan Cedex, France
E-mail: Jean-Michel.Morel@cmla.ens-cachan.fr

Professor F. Takens, Mathematisch Instituut,
Rijksuniversiteit Groningen, Postbus 800,
9700 AV Groningen, The Netherlands
E-mail: F.Takens@rug nl

Professor B. Teissier,
Institut Mathématique de Jussieu,
UMR 7586 du CNRS,
Équipe "Géométrie et Dynamique",
175 rue du Chevaleret,
75013 Paris, France
E-mail: teissier@math.jussieu fr

For the "Mathematical Biosciences Subseries" of LNM:

Professor P.K. Maini, Center for Mathematical Biology,
Mathematical Institute, 24-29 St Giles,
Oxford OX1 3LP, UK
E-mail: maini@maths.ox.ac.uk

Springer, Mathematics Editorial I, Tiergartenstr. 17,
69121 Heidelberg, Germany,
Tel.: +49 (6221) 487-8259
Fax: +49 (6221) 4876-8259
E-mail: lnm@springer.com