Thesis presented for the degree of
Dr. rer. nat

Potts models and stochastic dynamics

Daniel Meißner

December 2020

RUHR
UNIVERSITÄT
BOCHUM
RUB

Advisors:
Prof. Dr. Christof Külske
Prof. Dr. Peter Eichelsbacher

To Anika and Emil

Contents

1 Introduction

In this thesis we are investigating a time-evolved version of the symmetric mean-field Potts model. It is our goal to understand the critical parameters (both temperature and time) for which the model shows the *sequential Gibbs property*. This type of research has its origins in the discovery of non-Gibbs measures which arise from Gibbs measures under natural transformations. At that time, the Gibbs property of measures was merely taken for granted. In a founding paper van Enter, Fernández, and Sokal [56] considered these natural transformations and proved rigorously that in several cases the resulting measures are not Gibbs measures. A key concept in the understanding of these pathologies is the notion of a *hidden phase transition* which will be discussed later on.

The Potts model is a model from statistical physics that is a generalization of the famous Ising model. Despite its name "Ising model", the physicist Wilhelm Lenz invented it in 1920 [44] and asked his student Ernst Ising to study it in his PhD thesis [27]. Historically, the Ising model was considered a model for ferromagnetism. Today, it is one of the most-studied models in statistical physics and has also raised interesting problems in the mathematical theory of probability. This is most probably due to its simplicity and interesting behavior. At the time of Lenz and Ising, a magnet was imagined to be a collection of small magnetic moments that sit on some spatial structure like a lattice. Below a threshold temperature (the *critical temperature*) the magnetic moments in a ferromagnet have the tendency to align themselves with each other, even in the absence of an external magnetic field. In statistical physics this is known as a phase transition. The Ising model tries to mimic this phenomenon by modelling these magnetic moments as particles that can have the two states 1 and -1. These particles are then arranged on a spatial structure like a d-dimensional lattice, for example, and neighboring particles are said to interact in such a way that their

preference is to align themselves with their neighbors. In fact, in the first version of the Ising model (as it was considered by Ernst Ising) the lattice was even one-dimensional. Unfortunately, Ising was mislead by results from the one-dimensional model which does not show the expected phase transition, and falsely conjectured that also the models on the higher dimensional lattices would not show any phase transitions. Rudolph Peierls [48] later found this conjecture not to be true and instead proved that the Ising model does show a phase transition in two or more dimensions which led to the immense success of this model. Eight years after Peierls proof, in 1944, Lars Onsager [47] provided an exact solution of the Ising model. The Potts model, which was first considered by Renfrey Potts in 1951, generalizes the Ising model so that it allows the particles to have q different states (sometimes referred to as colors) where $q \geq 3$ is an integer. Like the Ising model, the Potts model favors configurations where neighboring particles have the same state.

Even before the Ising model, another model by Pierre-Ernest Weiss was a candidate for the modelling of ferromagnetism. Today, this model is known as the *Curie–Weiss model* or *mean-field Ising model* in the statistical physics and probability literature. In a mean-field model, the lattice is replaced by the complete graph, that is, every particle interacts with every other particle and favors the same state. Although this might seem to yield completely different models, mean-field models often show similar behavior to their lattice counterparts, at least in higher dimension. The one-dimensional Ising model, for example, does not show a phase transition but the Curie–Weiss model does. Already the two-dimensional Ising model, however, shows the same behavior as the Curie–Weiss model. There are even quantitative results concerning the critical temperatures of the two models [12]. To determine whether a mean-field model shows a phase transition, one can transform this into an optimization problem of a certain *potential* function, sometimes called *free energy*. This potential function depends on the parameters of the model and the model shows a phase transition if the global minimum of the function is degenerate, that is, if multiple global minimizers exist. Since it is usually hard to find sharp results for the critical parameters on the lattice, mean-field models are often considered solvable by this optimization approach. Our investigations of the mean field-Potts model, however, show that the task to understand the critical parameters at

which the minimizers of the potential change can be quite extensive.

There are two different methods to obtain the potential function. One is based on the theory of large deviations, the other is based on the so-called Hubbard-Stratonovič transformation. Essentially, the idea of the two approaches is very similar: They quantify on an exponential scale the concentration of probability mass on the different events. Whereas large deviation principles provide upper and lower bounds on this concentration in terms of a so-called *rate function*, the second approach uses the Hubbard-Stratonovič transformation to rewrite the probability distribution in such a way that *Laplace's method* [42] can be used. This means roughly the following: The main contribution to integrals of the form

$$\int e^{-nf(x)} \, \mathrm{d}x$$

for large n come from the neighborhoods of the global minimizers of f. The probability to observe a certain magnetization for a system with a large number n of particles therefore concentrates around the global minimizers of f. If the potential has multiple global minimizers, the physical interpretation is that different phases of the system (with different values for the magnetization) can coexist.

The potential function typically depends on a number of parameters. In the case of the mean-field Potts with three states these are: the inverse temperature β and a two-dimensional external field (u, v). The inverse temperature acts as a homogeneous coupling strength of the interaction of the particles and the external fields (u, v) model the bias of the system in favor of one or more states. Low temperature (high β) means that the particle interaction is very strong and they are likely in the same state, high temperature (low β) means that the particle interaction is weak. In order to understand the effect of the parameters u and v, it is instructive to refer to Figure 1.1. The bias introduced by the external fields (u, v) can be represented by a point in the unit simplex. Without any bias, we have $(u, v) = (0, 0)$ and we are in the center of the unit simplex. If the external fields are represented by a point near vertex 1, for example, it means that a-priori the system favors configurations which have a lot of particles in state 1. On the other hand, if we choose one of the midpoints of the simplex' edges, the system favors the two states of the adjacent vertices equally. Note that we can represent both the external

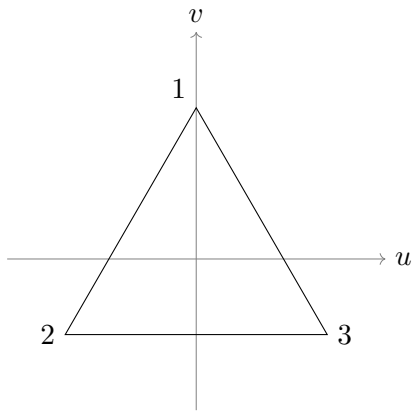

Figure 1.1: The unit simplex can be used as a representation for the bias that the external fields introduce for a Potts model with three states. For example, an external field represented by one of the simplex' vertices means that the system favors the respective state. If a midpoint of the simplex edges is considered, this means that the system favors the two states belonging to the adjacent vertices equally. In this way, we can describe the bias using the two coordinates u and v.

fields as well as the magnetization (the vector of the relative frequencies of the states $1, 2, 3$ in a configuration of particles) by a point in the unit simplex. In terms of *singularity theory* we are dealing with a function with a three-dimensional *control space* (the parameters β, u, v) and a two-dimensional *state space* (the magnetization (x, y)).

Singularity theory is concerned with the local change of a parameter-dependent smooth function of one or more variables near its critical points, that is, points of vanishing first derivative. This change is induced by a change of the parameters of the function. One important result of this theory is that the nature of these changes can be classified and that, surprisingly, the list of possible changes for r-dimensional parameters with $r \leq 4$ is finite and is known as *Thom's Seven Elementary Catastrophes*. The precise formulation of this classification is the content of Thom's theorem. This theorem will be discussed in more detail in the following chapter but let us mention that the equivalence relation imposed on the set of potential functions is based on parameter-dependent local diffeomorphisms acting on the domain and parameter-dependent

translations in the target space. This leads to a list of parameter-independent *germs* of functions which show a specific behavior when they are perturbed by parameter-dependent terms. These perturbations are referred to as *unfoldings* in singularity theory and the germs represent types of *singularities* which have figurative names like the *cusp*, the *butterfly* or the *elliptic umbilic*. For example, the polynomial prototype of the cusp singularity is given by $x^4 + ax^2 + bx$. Depending on the value of the parameters a and b this polynomial has either two minimizers or one minimizer. The case $b = 0$ is particularly easy to understand: Since $x^4 + ax^2 \to \infty$ as $|x| \to \infty$ and the number of inflection points is bounded by two, the number of minimizers is determined by the local behaviour around $x = 0$. This is where the parameter a comes into play. For $a < 0$ the function has a local maximum at $x = 0$ and thus two minimizers, for $a \geq 0$ the function has a minimum at $x = 0$ and because of the bound on the number of inflection points this minimum is unique.

For the investigation of phase transitions of mean-field models, which are connected to optimization problems of parameter-dependent potential functions, the results of this theory provide a useful tool. It is our interest to understand the change of the global minimizers of the potential as a function of the parameters. Combining the local analysis of singularity theory to understand how the minimizers change under a change of the parameters with a global analysis of all minimizers, we can identify the critical parameters for which the potential function has a degenerate global minimum.

There are various types of Gibbs properties of which we will mention two. The first type is known as the DLR-formalism named after Dobrushin, Lanford and Ruelle which is suited for a description of models that have spatial dependence structures like a lattice, for example [23]. The key idea is that a measure in infinite volume is a Gibbs measure if it can be reconstructed from its conditional probabilities of the particles in finite volume given the state of all particles on the outside of this volume. The second type is known as sequential Gibbs property and is applicable to mean-field models like the mean-field Potts model. This notion of Gibbsianness is much younger and was introduced by Häggström and Külske [25] because the DLR-formalism does not yield much information for mean-field models where there is no spatial structure as the particles sit on a complete graph. Infinite-volume limits of mean-field models

either lead to product measures (which are Gibbsian in the DLR-sense) or mixtures of product measures (which are non-Gibbsian in the DLR-sense). The sequential Gibbs property imitates a property of DLR-Gibbs measures called *quasilocality*. This property can be interpreted as a continuity property of the dependence of the conditional probabilities on the state of the system on the outside of the respective finite volume.

We will now describe the mathematical formalism associated with the concepts above and use them to define the symmetric mean-field Potts model. A system of $n \in \mathbb{N}$ particles in statistical physics is described by a probability distribution on the set of possible configurations Ω_n. The set of configurations for the system is an n-fold Cartesian product of the set E of states in which an individual particle can be in, that is, $\Omega_n = E^n$. For the symmetric mean-field Potts model with three states this set is $\Omega_n = \{1, 2, 3\}^n$ and the probability distribution on Ω_n is given by the *Boltzmann weights*

$$\mathbb{P}_n[(\omega_1, \ldots, \omega_n)] = \frac{e^{-\beta H_n(\omega_1, \ldots, \omega_n)}}{Z_n(\beta)}$$

where $\omega_i \in \{1, 2, 3\}$ and $i \in \{1, \ldots, n\}$. The *partition function* $Z_n(\beta)$ is the normalizing factor that makes the probabilities sum up to one, that is,

$$Z_n(\beta) = \sum_{(\omega_1, \ldots, \omega_n) \in \Omega_n} e^{-\beta H_n(\omega_1, \ldots, \omega_n)}.$$

The function $H_n \colon \Omega_n \to \mathbb{R}$ is called the *Hamiltonian* of the model. It can be interpreted as the interaction energy of the particles which leads to high probability for a configuration whose interaction energy is low and vice versa. We can now also understand how the inverse temperature β controls the interaction strength. Low values of β mean that the effects of the interaction energy is weakened, for example $\beta = 0$ would yield the uniform distribution on Ω_n, that is, all the particles are independent and all configurations are equally likely. In the case of the symmetric Potts model the Hamiltonian is defined as follows:

$$H_n(\omega_1, \ldots, \omega_n) = -\frac{1}{2n} \sum_{i,j=1}^{n} 1_{\sigma_i = \sigma_j}$$

This function clearly produces low values (and hence high Boltzmann weights) for configurations $(\sigma_1, \dots \sigma_n)$ where a lot of the particles have the same state.

A phase transition refers to an abrupt change of a physical quantity when external conditions are modified. For example, water shows a drastic change in volume when its temperature passes the boiling point. In statistical physics we want to explain the macroscopic properties of matter using microscopic models. The microscopic model is given in terms of a random variable that captures the full microscopic state of the system. An example for such a random variable in the case of the Curie–Weiss model is the random vector $(\omega_1, \dots, \omega_n) \in \{-1, 1\}^n$. This vector captures the magnetic moments of all n particles that the system consists of and thus captures the full microscopic state. The macroscopic property *magnetization* can be studied by analysing the limiting behavior of the observable $\frac{1}{n}\sum_{i=1}^{n} \omega_i$ as $n \to \infty$. The fact that the Curie–Weiss model shows a phase transition is reflected in the breakdown of the law of large numbers for $\beta > 1$,

$$\frac{1}{n}\sum_{i=1}^{n} \omega_i \to \frac{1}{2}(\delta_{-m(\beta)} + \delta_{m(\beta)}),$$

where $m(\beta)$ is some real number in the interval $(0, 1)$. If we approach the critical temperature of the Curie–Weiss model from below, that is, $\beta \searrow 1$, we find that $m(\beta) \to 0$ [19, Chapter 2]. This means that for a temperature above 1, the system shows no magnetization whereas below this temperature we find that it is magnetized. One should also note that for $\beta = 1$, that is, at the critical point, the law of large numbers applies but with an unusual type of central limit theorem [13]. This shows why the Curie–Weiss model is also interesting from other perspectives of probability theory.

A mean-field model is said to have the *sequential Gibbs property* if certain conditional probabilities have a well-defined limit as the number of particles tends to infinity. In the language of statistical physics this limit is called the *thermodynamic limit*. To be more precise, we consider the finite-volume conditional probability $\mathbb{P}_n[\sigma_1 | \sigma_2, \dots, \sigma_n]$ of the first particle σ_1 given a fixed configuration $(\sigma_2, \dots, \sigma_n)$ of the remaining particles. For the sequential Gibbs property we demand the existence of

$\gamma(\sigma_1|\alpha) := \lim_{n\to\infty} \mathbb{P}_n[\sigma_1|\sigma_{n,2},\ldots,\sigma_{n,n}]$ where $(\sigma_1,\sigma_{n,2},\ldots,\sigma_{n,n}) \in \Omega_n$ is a sequence of configurations for an increasing number of particles such that the empirical measures

$$\frac{1}{n-1}\sum_{k=2}^{n} \delta_{\sigma_{n,k}}$$

have the limit α for as $n \to \infty$. If the limit $\gamma(\sigma_1|\alpha)$ exists, we call α *good*. The mean-field model is called sequentially Gibbs if every probability measure on E is good.

The symmetric mean-field Potts model is sequentially Gibbs in this sense and in view of the results of van Enter, Fernández, and Sokal [56] a natural question to ask is whether this property is preserved for our time-evolved version of the mean-field Potts model. A key concept in this investigation is the notion of a *hidden phase transition*. The reason for the loss of the continuity of the conditional probabilities in the infinite-volume limit lies in a discontinuity in another system of particles. This system captures the state of the particles at time zero but conditioned to end in a certain configuration at a later point $t > 0$ in time. It is known as the *constrained first-layer model*.

This thesis is split in three parts. The first part explains the methods that we use in the other two parts of this thesis. The methods derive from two different fields of mathematics: probability theory and singularity theory. From probability theory we use the theory of large deviations and another method to establish limit laws sometimes referred to as *Hubbard-Stratonovič approach*. Singularity theory allows us to study the parameter-dependent functions that appear in the probabilistic methods. Two important concepts of singularity theory are the *catastrophe map* and the *bifurcation set*. By means of the catastrophe map we are able to analyse the bifurcation set which in turn helps us to understand the critical transition lines (Maxwell sets) of the static and time-evolved Potts model. This part is found in Chapter 2.

The second part is a detailed analysis of the mean-field Potts model in vector-valued external field contained in Chapter 3 which is also useful for the analysis of the time-evolved model. We are analysing the *free energy* which is a parameter-dependent real-valued function on the set of probability measures on $\{1,2,3\}$, that is, on the two-dimensional unit simplex Δ^2 without boundary. This function plays a

role in the large deviation principle that is satisfied by the mean-field Potts distributions. If this function has a degenerate global minimum, that is, if it has multiple global minimizers, the model is said to undergo a phase transition because the probability mass concentrates around multiple points in Δ^2. These points are interpreted as the possible phases of the system. The mean-field Potts model can at most have a coexistence of four phases, three ordered phases and one disordered phase corresponding to the global minimizers of the free energy. Depending on the external field, which is equivalently described by an a-priori measure α on $\{1, 2, 3\}$, the number of global minimizers varies between one and four. This behavior is analysed via a discussion of the possible two-dimensional sections of the Maxwell set for fixed inverse temperatures β. The Maxwell set is a subset of the three-dimensional parameter space $(0, \infty) \times \Delta^2$, consisting of inverse temperature β and external field α, which contains those parameters (β, α) such that the free energy has at least two global minimizers. Moreover, with our method we also gain insight into the metastable, that is, local minima of the free energy. This results in a decomposition of the parameter space into regions where the number of local and global minima is constant.

In the last part in Chapter 4 we study a natural time-evolution of the symmetric mean-field Potts model. This time-evolution is a symmetric spin-flip dynamics that is independent over the different particles. This means that the future state of a particle at time $t > 0$ given a configuration of the system at time $t = 0$ depends only on the state of the same particle in the past. In other words, each particle evolves independently of the other particles. The physical interpretation is that the particles are heated in a bath of infinite temperature which is a special case of *Glauber dynamics*. We are then interested in the sequential Gibbs property of the time-evolved measures. Using the two-layer approach we are able to reduce this question to the study of phase transitions of the first-layer model constrained to have empirical distribution α at time t. In the analysis of the parameter-dependent potential function that arises in the Hubbard-Stratonovič approach we are able to use the knowledge of the static model. We show that in a mid-range temperature regime ($\beta < 3$) the transitions of the static model re-appear where α plays the role of the external fields and the time parameter t allows for a perturbation of the bifurcation set slices we observe in the static case. However, unlike

the static case where the external fields are constrained to the interior of the unit simplex, in the time-evolved model we find various exit scenarios of the bifurcation set slices. More precisely, some parts of the bifurcation set leaves the unit simplex and therefore the corresponding bad measures α are not visible. Complementing the analysis with a combination of exact symbolic computation and numerics, we are able to provide a full analysis of this mid-range temperature regime. In particular, we describe the time-evolution of the bad measures for different subregions in the regime $\beta < 3$. The behavior of this model shows a number of interesting properties, first and foremost the recovery of the sequential Gibbs property. Additionally, unlike the Curie–Weiss model, the non-Gibbs temperature, which is defined to be the temperature above which no transitions between Gibbs and non-Gibbs occur, does not coincide with the critical temperature of the static model. Furthermore, we are able to fully analyse the trajectories of bad empirical measures and we find essentially five β-regimes with different sequences of bad measures appearing in the time-evolution. For every $\beta < 3$ the bad empirical measures first appear (if they appear at all) as the midpoints of the simplex edges and form three straight lines moving towards the center of the simplex. With increasing β the time-evolution of the set of bad measures becomes more involved. The three lines branch off on one end to two curves of fully asymmetric bad measures. Later these Y-shaped sets become connected and depending on the β-regime the partially symmetric part of these sets either leaves the simplex before the sets connect or after they have connected. For a full overview see Figure 4.3. In the regime $\beta \geq 3$ the analogy with the static model breaks down and new types of transitions appear which is left to future research.

2 Gibbs–non-Gibbs transitions and singularity theory

In this chapter we are establishing the connection between the sequential Gibbs property of our time-evolved symmetric mean-field Potts model and the ideas and terms from singularity theory. In the first part we will describe the *constrained first-layer model* that emerges in the *two-layer approach* and present two methods for the analysis of the concentration behavior. In the second part we give the definitions for the basic objects in singularity theory. We also cover one major result of the theory: Thom's classification theorem.

2.1 From probability theory to singularity theory

The mean-field Potts model describes a system of a finite number n of particles that can have q different colors. Therefore the *configuration space* Ω_n is the set of all tuples $(\omega_1, \ldots, \omega_n)$ with $\omega_i \in \{1, \ldots, q\}$. Note that because these spaces are finite, we can describe the probability measures on Ω_n by a *probability vector*. In the Chapter 3 of this thesis, we are studying the mean-field Potts model with $q = 3$ states in a vector-valued external field. This external field is represented by a point α in the (interior of the) unit simplex

$$\Delta^{q-1} = \left\{ (\alpha_1, \ldots, \alpha_q) \;\middle|\; \sum_{b=1}^{q} \alpha_b = 1, \forall b \in \{1, \ldots, q\} : \alpha_b > 0 \right\}.$$

Definition 1. *The* mean-field Potts model at inverse temperature $\beta > 0$ and external field $\alpha \in \Delta^2$ *is the sequence* $(\mu_{n,\beta,\alpha})_{n\in\mathbb{N}}$ *where* $\mu_{n,\beta,\alpha}$ *is the probability measure on* Ω_n *given by*

$$\mu_{n,\beta,\alpha}(\omega_1, \ldots, \omega_n) = \frac{e^{\frac{\beta}{2n}\sum_{i,j=1}^{n} 1_{\omega_i=\omega_j}}}{Z_n(\beta, \alpha)} \prod_{i=1}^{n} \alpha_{\omega_i} \tag{2.1}$$

where the partition function *is defined as*

$$Z_n(\beta,\alpha) = \sum_{(\omega_1,\dots,\omega_n)\in\Omega_n} e^{\frac{\beta}{2n}\sum_{i,j=1}^n 1_{\omega_i=\omega_j}} \prod_{i=1}^n \alpha_{\omega_i}. \tag{2.2}$$

The static mean-field Potts model in external field is a first step for the understanding of the time-evolved version of the model as we will see in Chapter 4. Although our analysis in this chapter is for the case $q = 3$, we will define the dynamics for an arbitrary integer $q \geq 2$. The time-evolution will be given in terms of a Markov chain that is constructed out of independent Markov chains which govern the time-evolution of the individual particles. Therefore, we are given *transition probabilities* $p_t(a, b)$ that describe the probability that a particle in state $a \in \{1, \dots, q\}$ is in state $b \in \{1, \dots, q\}$ at time $t > 0$. The particles undergo a *symmetric spin-flip dynamics*, that is, they change their color to i at a rate $r_i = 1$ for all $i \in \{1, \dots, q\}$. This leads to the transition probabilities

$$p_t(a,b) = \frac{e^{g_t 1_{a=b}}}{e^{g_t} + q - 1}$$

where $g_t := \log(1 + 2e^{-3t}) - \log(1 - e^{-3t})$. We will now define the time-evolved measures $\mu_{n,\beta,t}$.

Definition 2. *The* time-evolved mean-field Potts model at time $t > 0$ with initial inverse temperature β *is given by the sequence* $(\mu_{n,\beta,t})_{n\in\mathbb{N}}$ *of probability measures on* Ω_n *given by*

$$\mu_{n,\beta,t}(\eta_1,\dots,\eta_n) = \sum_{(\omega_1,\dots,\omega_n)\in\Omega_n} \frac{e^{\frac{\beta}{2n}\sum_{i,j=1}^n 1_{\omega_i=\omega_j}}}{q^n Z_n(\beta,\alpha_0)} \prod_{i=1}^n p_t(\omega_i,\eta_i) \tag{2.3}$$

where α_0 *equals the uniform distribution on* $\{1, \dots, q\}$.

These measures represent a time-evolution of the *symmetric* mean-field Potts model since we are using the uniform distribution α_0 as an a-priori measure which equivalently describes the absence of an external field. This means that initially there is no bias towards any particular color.

We are then interested in the parameters (β, t) for which the above time-evolved mean-field Potts model shows the *sequential Gibbs property*. This definition essentially requires that we can define an infinite-volume limit of the conditional probabilities in finite volume.

Definition 3. *We say that $\alpha \in \Delta^{q-1}$ is a* good point *for the time-evolved mean-field Potts model $(\mu_{n,\beta,t})_{n\in\mathbb{N}}$ if and only if the limit*

$$\gamma_{\beta,t}(\eta_1|\alpha) := \lim_{n\to\infty} \mu_{n,\beta,t}(\eta_1|\eta_{n,2},\dots,\eta_{n,n}) \tag{2.4}$$

exists for every family $\eta_{n,k} \in \{1,\dots,q\}$ with $n \geq 2$ and $k \in \{2,\dots n\}$ such that

$$\lim_{n\to\infty} \frac{1}{n-1} \sum_{k=2}^{n} \eta_{n,k} = \alpha. \tag{2.5}$$

Otherwise, we call α a bad point. *We say that the time-evolved mean-field Potts model $(\mu_{n,\beta,t})_{n\in\mathbb{N}}$ is* sequentially Gibbs *if all α in the unit simplex Δ^{q-1} are good points.*

One should note that the original definition of the sequential Gibbs property involved an additional condition [25]. This condition asked that the function $\alpha \mapsto \gamma_{\beta,t}(\eta_1|\alpha)$, which is defined in the limit process, is continuous. However, this requirement is already implied by the existence of the limit itself [26].

We will now explain the *two-layer approach*, which was used, for example, by van Enter, Fernández, and Sokal [56] to prove the non-Gibbsianness of a measure resulting from a block-spin transformation of a Gibbsian measure. However, they considered lattice models and the DLR-Gibbs property. Nevertheless, we can use the same idea: We consider a two-layer system where the first layer consists of the configuration of the model at time zero and the second layer consists of the configuration at time $t > 0$. The loss of the sequential Gibbs property at time $t > 0$ is due to a discontinuity in an internal system. This internal system is the *constrained first-layer model* and it is *hidden* beneath the second layer. This is why we refer to a phase transition in the constrained first-layer model as a *hidden phase transition*. We will now be more precise and define the constrained first layer model. Then we will describe the relation of this model to the conditional probabilities in finite volume.

Definition 4. *The* constrained first-layer model *of the time-evolved mean-field Potts model is a sequence $(\bar{\mu}_n)_{n\in\mathbb{N}}$ of mappings from $\Omega_{[2,n]}$ to*

the space of probability measures on $\Omega_{[2,n]}$ *given by*

$$\bar{\mu}_n[\eta_2,\dots,\eta_n](\omega_2,\dots,\omega_n) = \frac{e^{\frac{\beta}{2n}\sum_{i,j=2}^{n} 1_{\omega_i=\omega_j}} \prod_{i=2}^{n} p_t(\omega_i,\eta_i)}{\sum_{(\tilde{\omega}_2,\dots,\tilde{\omega}_n)\in\Omega_{[2,n]}} e^{\frac{\beta}{2n}\sum_{i,j=2}^{n} 1_{\tilde{\omega}_i=\tilde{\omega}_j}} \prod_{i=2}^{n} p_t(\tilde{\omega}_i,\eta_i)} \tag{2.6}$$

where $\Omega_{[2,n]} = \{(\omega_2,\dots,\omega_n) \mid \forall i \in \{2,\dots,n\} : \omega_i \in \{1,\dots,q\}\}$.

The goal of this section is to connect the probabilistic notion of *sequentially Gibbs* to a parameter-dependent optimization problem that we can analyse using methods from singularity theory presented in the following section. This is done in two steps. In the first step we represent the finite-volume conditional probabilities which appear in Formula (2.4) in terms of an expected value with respect to the constrained first-layer model. In the second step we analyse the possible phase transitions of the constrained first-layer model, which is a model with a four-dimensional parameter. Because of the identity

$$\mu_{n,\beta,t}(\eta_1|\eta_2,\dots,\eta_n) = \left\{ \sum_{\tilde{\eta}_1} \frac{\mu_{n,\beta,t}(\tilde{\eta}_1,\dots,\eta_n)}{\mu_{n,\beta,t}(\eta_1,\dots,\eta_n)} \right\}^{-1},$$

it is convenient to study the quotients of the finite-volume time-evolved measures.

Lemma 5. *The quotients of the time-evolved measures have the representation*

$$\frac{\mu_{n,\beta,t}(\tilde{\eta}_1,\dots,\eta_n)}{\mu_{n,\beta,t}(\eta_1,\dots,\eta_n)} = \frac{\int f_n^{\tilde{\eta}_1}\, \mathrm{d}\bar{\mu}_n[\eta_2,\dots,\eta_n]}{\int f_n^{\eta_1}\, \mathrm{d}\bar{\mu}_n[\eta_2,\dots,\eta_n]} \tag{2.7}$$

where the observable $f_n^{\eta_1}$ *is given by*

$$f_n^{\eta_1}(\omega_2,\dots,\omega_n) = \sum_a \exp\left(\frac{\beta}{n}\sum_{i=2}^{n} 1_{\omega_i=a}\right) p_t(a,\eta_1). \tag{2.8}$$

Proof. The proof is done by separating the terms containing η_1 from the terms containing $\eta_2, \ldots, \eta_n$. First observe that

$$\sum_{i,j=1}^{n} 1_{\tilde{\omega}_i=\tilde{\omega}_j} = \sum_{i,j=2}^{n} 1_{\tilde{\omega}_i=\tilde{\omega}_j} + 2\sum_{i=2}^{n} 1_{\tilde{\omega}_i=\tilde{\omega}_1} + 1. \tag{2.9}$$

Then we find

$$\frac{\mu_n(\tilde{\eta}_1 \ldots, \eta_n)}{\mu_n(\eta_1 \ldots, \eta_n)} = \frac{\sum_{\tilde{\omega}_2,\ldots,\tilde{\omega}_n} f_n^{\tilde{\eta}_1} \exp\left(\frac{\beta}{2n} \sum_{i,j=2}^{n} 1_{\tilde{\omega}_i=\tilde{\omega}_j}\right) \prod_{i=2}^{n} p_t(\tilde{\omega}_i, \eta_i)}{\sum_{\tilde{\omega}_2,\ldots,\tilde{\omega}_n} f_n^{\eta_1} \exp\left(\frac{\beta}{2n} \sum_{i,j=2}^{n} 1_{\tilde{\omega}_i=\tilde{\omega}_j}\right) \prod_{i=2}^{n} p_t(\tilde{\omega}_i, \eta_i)} \tag{2.10}$$

If we introduce the normalizing factor

$$\sum_{\tilde{\omega}_2,\ldots,\tilde{\omega}_n} \exp\left(\frac{\beta}{2n} \sum_{i,j=2}^{n} 1_{\tilde{\omega}_i=\tilde{\omega}_j}\right) \prod_{i=2}^{n} p_t(\tilde{\omega}_i, \eta_i) \tag{2.11}$$

into the numerator and denominator, we arrive at the representation (2.7). □

Note that we can write $f_n^{\eta_1}(\omega_2, \ldots, \omega_n) = \tilde{f}_n^{\eta_1}\left(\frac{1}{n-1} \sum_{i=2}^{n} \delta_{\omega_i}\right)$ where

$$\tilde{f}_n^{\eta_1}(\nu) = \sum_a e^{\beta \frac{n-1}{n} \nu_a} p_t(a, \eta_1)$$

for $\nu \in \Delta^{q-1}$ and that $\tilde{f}_n^{\eta_1}$ converges uniformly as $n \to \infty$. Therefore, it suffices to study the weak convergence of the random variable $\frac{1}{n-1} \sum_{i=2}^{n} \delta_{\omega_i}$ where $\omega_2, \ldots, \omega_n$ have joint law $\mu_n[\eta_{n,2}, \ldots, \eta_{n,n}]$ and where $\eta_{n,k}$ fulfills (2.5). We can analyse this convergence with the help of two different methods: Large deviation principles or the Hubbard-Stratonovič (HS) approach. We will discuss the necessary steps for the large deviations approach here and why we are not using it for the time-evolved mean-field Potts model. In Chapter 4 we will use the HS approach instead.

A *large deviations principle* consists of an upper and a lower bound for a sequence of probability measures. These bounds play an important role in the analysis of rare events, that is, exponentially small probabilities. Consider a sequence of probability spaces $(\Omega, \mathcal{B}, \mathbb{P}_n)$ where Ω is a Polish space and $\mathcal{B}$ is the Borel σ-field.

Definition 6. *The sequence* $(\mathbb{P}_n)_{n\in\mathbb{N}}$ *is said to fulfill a* large deviations principle (LDP) *with rate function* $I\colon \Omega \to [0,\infty]$ *if all of the following conditions are fulfilled:*

(1) I *has compact sub-level sets and* $I(x) < \infty$ *for some* $x \in \Omega$.

(2) For all closed sets $C \in \mathcal{B}$ *we have*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(C) \leq - \inf_{x\in C} I(x).$$

(3) For all open sets $O \in \mathcal{B}$ *we have*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_n(O) \geq - \inf_{x\in O} I(x).$$

One should note that some authors replace the assumption that I has compact sub-level sets by the assumption that the sub-level sets are relatively closed in Ω and call the rate function a *good rate function* if the compactness-assumption also applies. Let us make a short remark about the connection between an LDP and the weak law of large numbers. For every global minimizer x_0 of I we have $I(x_0) = 0$: With $C = \Omega$ it follows from point 2 that $\inf_{x\in\Omega} I(x) \leq 0$. Because $I(x) \geq 0$ for $x \in \Omega$ it follows that $I(x_0) = 0$. Note that there is always at least one such x_0 because the sub-level sets of I are compact (point 1 in the definition). Suppose that x_0 is unique, then the Portemanteau Theorem implies that $\mathbb{P}_n$ converges weakly to the Dirac measure δ_{x_0}. Therefore if the constrained first-layer models fulfills an LDP for a rate function with a unique global minimizer, it follows that it fulfills the weak law of large numbers. In other words, because of the representation (**??**) of the finite-volume conditional probabilities the α under consideration is a good point. This is why we are interested in the global minimizers of the rate function.

In order to derive a large deviations principle for the constrained first-layer model, we will use two basic theorems: Sanov's theorem and Varadhan's lemma/Tilted LDP. This allows us to compute the rate function for the LDP. As we will see, however, due to the structure of the

measure, this rate function will live on a $q(q-1)$-dimensional state space. So even in the case $q = 3$ a useful visualization of the rate function is not immediate. The Hubbard-Stratonovich approach described in Chapter 4 yields an optimization problem for a potential function on $\mathbb{R}^q$.

Theorem 7 (Varadhan's lemma)**.** *Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures that satisfies the LDP with rate function $I\colon \Omega \to [0,\infty]$ and let $\phi\colon \Omega \to \mathbb{R}$ be any continuous function. If*

$$\lim_{M\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log \int_{\phi\geq M} e^{n\phi(x)} \mu_n(\mathrm{d}x) = -\infty, \tag{2.12}$$

then

$$\lim_{n\to\infty} \frac{1}{n} \log \int e^{n\phi(x)} \mu_n(\mathrm{d}x) = \sup_{x\in\Omega}(\phi(x) - I(x)) \tag{2.13}$$

As a consequence of Varadhan's lemma, we can prove the *Tilted LDP*. This theorem states a large deviations principle for a family of measures that is given in terms of exponential densities with respect to another family fulfilling a large deviations principle. This is exactly the structure of the constrained first-layer model.

Theorem 8 (Tilted LDP)**.** *Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures that satisfies the LDP with rate function $I\colon \Omega \to [0,\infty]$ and let $\phi\colon \Omega \to \mathbb{R}$ be any continuous function which fulfills* (2.12)*. Then the sequence of probability measures $\tilde{\mu}_n$ defined via*

$$\int f \,\mathrm{d}\tilde{\mu}_n := \frac{\int f e^{n\phi(x)} \mu_n(\mathrm{d}x)}{\int e^{n\phi(x)} \mu_n(\mathrm{d}x)} \tag{2.14}$$

for any $f \in C_b(\Omega)$ fulfills the LDP with the rate function J given by

$$J(x) = \sup_{y\in\Omega}(\phi(y) - I(y)) - (\phi(x) - I(x)) \tag{2.15}$$

for any $x \in \Omega$.

For the proofs of the above theorems we refer to [9] and [33]. Using the Tilted LDP theorem, we can prove the an LDP for the constrained first-layer model. Define the vector $\mathbf{L}_n = (L_b^n)_{b=1}^q \in (\Delta^{q-1})^q$ of empirical measures on the b-like sites via

$$L_b^n = \frac{1}{|\Lambda_n(b)|} \sum_{i\in\Lambda_n(b)} \delta_{\omega_i} \tag{2.16}$$

where $\Lambda_n(b) = \{i \in \{2, \dots, n\} \mid \eta_{n,i} = b\}$ are the so-called *b-like sites*. Note that the constrained first-layer model has a density like in Theorem 8 with respect to the measures $\omega_i \mapsto \prod_{i=2}^n p_t(\omega_i, \eta_{n,i})$. In a first step, we prove a Sanov-like theorem for the vector $\mathbf{L}_n$ with respect to these measures. Because of the product structure, the components of the vector $\mathbf{L}_n$ are independent. Therefore we can use Sanov's theorem [9, Theorem 2.1.10] on each component because $\eta_{n,i} = b$ for every ω_i with $i \in \Lambda_n(b)$ and therefore the ω_i are independent and identically distributed. This extends to an LDP of the product space. In the second step, we can use the Tilted LDP to prove the following LDP for the constrained first-layer model.

Theorem 9. *Suppose the family $\eta_{n,k} \in \{1, \dots, q\}$ with $n \geq 2$ and $2 \leq k \leq n$ fulfills* (2.5)*. Then the sequence of image measures $\bar{\mu}_n[\eta_{n,2}, \dots, \eta_{n,n}] \circ \mathbf{L}_n^{-1}$ fulfills the LDP with rate function $\Phi_{\alpha,\beta,t} - \inf \Phi_{\alpha,\beta,t}$ where*

$$\Phi_{\alpha,\beta,t}(\nu) = -\frac{1}{2}\beta \sum_a \Big(\sum_b \alpha_b \nu_{b,a} \Big)^2 - g_t \sum_b \alpha_b \nu_{b,b} + \sum_{a,b} \alpha_b \nu_{b,a} \log \nu_{b,a} \tag{2.17}$$

for $\nu \in (\Delta^{q-1})^q$.

So in order to analyse the global minimizers of the large deviations rate function $\Phi_{\alpha,\beta,t}$, let us compute the derivative. Let $(\rho_{b,a})_{b,a=1}^q$ be any element of the tangent space of $(\Delta^{q-1})^q$. Then we find

$$\begin{aligned} \Phi'_{\alpha,\beta,t}(\nu)\rho &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \Phi(\nu + t\rho) \\ &= \sum_{b,a} \alpha_b \rho_{b,a} \Big(\log \nu_{b,a} - \beta \sum_c \alpha_c \nu_{c,a} - g_t 1_{b=a} \Big). \end{aligned} \tag{2.18}$$

Since $\sum_a \rho_{b,a} = 0$ for every $b \in \{1, \dots q\}$, the equation $\Phi'_{\alpha,\beta,t}(\nu) = 0$ is equivalent to

$$\log \nu_{b,a} - \beta \sum_c \alpha_c \nu_{c,a} - g_t 1_{a=b} = C_b \tag{2.19}$$

for some constant vector $(C_b)_{b=1}^q$. Furthermore, we are able to determine this constant because of the constraint $\sum_a \nu_{b,a} = 1$. Define the map $\Gamma \colon \mathbb{R}^q \times (0, \infty) \to \mathrm{Mat}(q, \mathbb{R})$ via its components

$$\Gamma_{b,a}(M, t) = \frac{e^{M_a + g_t 1_{b=a}}}{\sum_c e^{M_c + g_t 1_{b=c}}} \tag{2.20}$$

for $b, a \in \{1, \ldots, q\}$. Then we can rewrite (2.19) as $\nu_{b,a} = \Gamma_{b,a}(\beta m, t)$ where m is given via its components

$$m_a = \sum_{a=1}^{q} \alpha_b \nu_{b,a}.$$

Multiplying both sides of $\nu_{b,a} = \Gamma_{b,a}(\beta m, t)$ by α_b, summing oven $b \in \{1, \ldots, q\}$, we arrive at the so-called *mean-field equation* $m = \alpha\Gamma(\beta m, t)$, which determines the critical points of $\Phi_{\alpha,\beta,t}$. Note that $\alpha\Gamma(\beta m, t)$ has to be interpreted as a row–matrix product. Because this equation is a fixed-point equation, we can use the contraction principle to understand at least some part of the *Gibbs region*

$$\mathcal{G} = \{(\beta, t) \mid \mu_{n,\beta,t} \text{ is sequentially Gibbs}\},$$

that is, we obtain a lower bound on the Gibbs region in the following sense.

Theorem 10. *For every $\beta > 0$ there is a $t_*(\beta) > 0$ such that the set of (β, t) with $t < t_*(\beta)$ is contained in the Gibbs region $\mathcal{G}$.*

An explicit formula for $t_*(\beta)$ is also available. For the proof of the theorem we use the contraction principle for the map $m \mapsto \alpha\Gamma(\beta m, t)$.

Proof. Suppose m is a solution of the mean-field equation. Then m is a probability vector. Therefore to conclude that $\Phi_{\alpha,\beta,t}$ has a unique global minimizers, we can use the contraction principle for the restriction of the map $m \mapsto \alpha\Gamma(\beta m, t)$ to the unit simplex Δ^{q-1} of probability measures. This set forms a complete metric space with the metric induced by the 1-norm on $\mathbb{R}^q$ and $m \mapsto \alpha\Gamma(\beta m, t)$ is a contraction on this metric space if $L(q, \beta, t) < 1$ as given in the next lemma. For fixed β this inequality can be solved for t and thus defines a $t_*(\beta)$. □

Lemma 11. *Consider the $(q-1)$-dimensional unit simplex Δ^{q-1} in $\mathbb{R}^q$ with the metric induced by the 1-norm. For all probability measures α the map $m \mapsto \alpha\Gamma(\beta m, t)$ on Δ^{q-1} is Lipschitz-continuous with Lipschitz constant*

$$L(q, \beta, t) = \beta \frac{(e^\beta + q - 2)^2 + 2e^{g_t}(e^\beta + q - 2)}{(e^{g_t} + e^\beta + q - 2)^2} \tag{2.21}$$

Proof. Let m, m' be in Δ^{q-1}. Since α is a probability measure, we obtain

$$\|\alpha\Gamma(\beta m,t)-\alpha\Gamma(\beta m',t)\| \le \|\Gamma(\beta m,t)-\Gamma(\beta m',t)\|$$

where we use the associated operator norm for α considered as a linear map from $\mathrm{Mat}(q,\mathbb{R}^q)$ to $\mathbb{R}^q$. By the mean value theorem it follows that a Lipschitz constant is given by $\beta \sup_m \|D_1\Gamma(\beta m,t)\|$. Calculating the norm of the derivative yields that

$$\begin{aligned}\|D_1\Gamma(\beta m,t)\| &= \sup_b \sum_a \sup_c \left|\frac{\partial\Gamma_{b,a}}{\partial M_c}\right| \\ &= \sup_b \sum_a \sup_c \Big|\delta_{a,c}-\Gamma_{b,c}(\beta m,t)\Big|\Gamma_{b,a}(\beta m,t) \\ &= \sup_b \sum_a (1-\Gamma_{b,a}(\beta m,t))\Gamma_{b,a}(\beta m,t) \\ &= \sup_b \left(1-\sum_a \Gamma_{b,a}^2(\beta m,t)\right)\end{aligned} \tag{2.22}$$

because $\sup_c |\delta_{a,c}-\Gamma_{b,c}(\beta m,t)| = 1-\Gamma_{b,a}(\beta m,t)$. Using the estimate

$$\begin{aligned}\sum_a (1-\Gamma_{b,a})\Gamma_{b,a} &= (1-\Gamma_{b,b})\Gamma_{b,b}+\sum_{a\neq b}(1-\Gamma_{b,a})\Gamma_{b,a} \\ &\le (1-\Gamma_{b,b})\Gamma_{b,b}+(1-\Gamma_{b,b}) \\ &= 1-\Gamma_{b,b}^2\end{aligned} \tag{2.23}$$

we can give an upper bound for the supremum of the norm of the derivative:

$$\sup_m \|D_1\Gamma(\beta m,t)\| \le 1-(\inf_{m,b}\Gamma_{b,b}(\beta m,t))^2 \tag{2.24}$$

Furthermore, because $\nu \mapsto \sum_{c\neq b} e^{\beta\nu_c}$ is strictly convex and thus assumes its maximum only in the extremal points of the simplex, we have the following lower bound on

$$\begin{aligned}\Gamma_{b,b}(\beta m,t) &= \frac{e^{\beta m_b+g_t}}{e^{\beta m_b+g_t}+\sum_{c\neq b}e^{\beta m_c}} = \frac{e^{g_t}}{e^{g_t}+\sum_{c\neq b}e^{\beta(m_c-m_b)}} \\ &\ge \frac{e^{g_t}}{e^{g_t}+\sum_{c\neq b}e^{\beta m_c}} \ge \frac{e^{g_t}}{e^{g_t}+e^{\beta}+q-2}\end{aligned} \tag{2.25}$$

Finally, we arrive at

$$\begin{aligned}\beta \sup_m \|D_1\Gamma(\beta m, t)\| &\leq \beta\left(1 - \frac{e^{2g_t}}{(e^{g_t} + e^{\beta} + q - 2)^2}\right) \\ &= \beta\frac{(e^{\beta} + q - 2)^2 + 2e^{g_t}(e^{\beta} + q - 2)}{(e^{g_t} + e^{\beta} + q - 2)^2}\end{aligned} \tag{2.26}$$

□

This lower bound on the Gibbs region $\mathcal{G}$ shows that our model is sequentially Gibbsian in a neighborhood of $t = 0$, and this property is sometimes called *short-time Gibbsianness*.

2.2 Ideas from singularity theory

Singularity theory is concerned with the study and classification of degenerate critical points, that is, points where the first derivative of a potential function vanishes and the Hessian matrix is degenerate. Let us start with the simplest example of a singularity: The fold.

Consider the family of functions $(V_a)_{a\in\mathbb{R}}$ with $V_a(x) = x^3 + ax$ for $x \in \mathbb{R}$. We want to study how the set of critical points changes if we vary $a \in \mathbb{R}$. The set of critical points of the potential function V_a is given by

$$\{x \in \mathbb{R} \mid 3x^2 + a = 0\}.$$

Now because it is a simple example, we see immediately that at $a = 0$ the structure of critical points changes: As long as a is positive, the quadratic equation has no real solutions so that the set of critical points is empty. For $a = 0$ we have a single critical point at the origin and for $a < 0$ we have two critical points. Using the second derivative $V_a''(x) = 6x$, we see that V_a has local minimum at $x = \sqrt{-\frac{a}{3}}$ and a local maximum $x = -\sqrt{-\frac{a}{3}}$. For $a = 0$ the potential function V_a has a single critical point at $x = 0$ that is a saddle point.

We will now analyse this example with concepts from singularity theory. More generally, for a family $(V_a)_{a\in\mathbb{R}^m}$ of potential functions $V_a \colon \mathbb{R}^n \to \mathbb{R}$

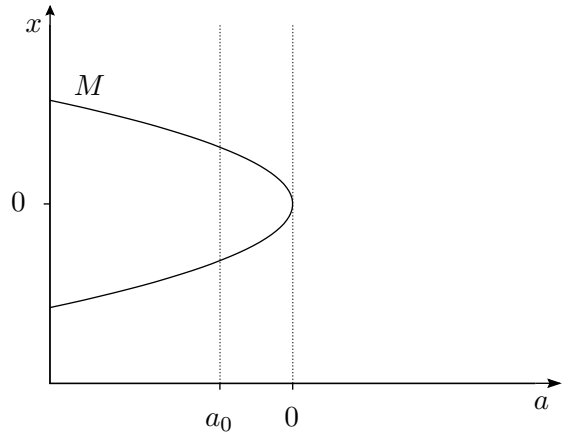

Figure 2.1: The catastrophe manifold M of the fold. The change in the number of critical points is clearly visible. For $a < 0$ there are two critical points whereas for $a > 0$ there are no critical points. At $a = 0$ the tangent vector along the curve M is vertical which characterizes a change in the set of critical points.

we call the space $\mathbb{R}^m$ the *control space* and the space $\mathbb{R}^n$ the *state space*. The *catastrophe manifold* for this family is defined as

$$M = \{(a, x) \in \mathbb{R}^{m+n} \mid \nabla V_a(x) = 0\}.$$

In contrast to the set of critical points we considered above, in the *catastrophe manifold* the parameter a is not fixed. Instead, the catastrophe manifold consists of pairs (a, x) such that x is a critical point of V_a. Since V_a is a smooth function, this set is indeed a smooth manifold. In order to understand the boundary of the parameter regions in which the number of critical points is constant, we can also use the following geometric description. Figure 2.1 shows the catastrophe manifold of the fold as a subset of the Cartesian product $\mathbb{R}^m \times \mathbb{R}^n$ of the parameter space $\mathbb{R}^m$ and the state space $\mathbb{R}^n$. In this case, $n = m = 1$ and we can clearly see that the line $a = 0$ separates $\mathbb{R}^2$ into two half-spaces. In the left half-space, we find two intersections of the line $a = a_0$ with the catastrophe manifold for each $a_0 < 0$. However, in the right half-space there are no such intersections. If we look at the catastrophe manifold in Figure 2.1, we see that, at the origin, it "folds under itself". We can characterize this

behavior as follows: The projection $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m, (a, x) \mapsto a$ restricted to the catastrophe manifold M, which is known as the *catastrophe map* χ, has a vertical tangent vector at $(0, 0)$. This leads us to the definition of the *bifurcation set*. This set is a subset of the control space $\mathbb{R}^m$ and is defined as follows:

$$B = \{a \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n : d\chi(a, x) = 0\}$$

Here, $d\chi$ is the differential of the catastrophe map $\chi \colon M \to \mathbb{R}^m$ with $\chi(a, x) = a$. Using $x \mapsto (-3x^2, x)$ as a chart for the catastrophe manifold M we consider χ in local coordinates: $x \mapsto -3x^2$. So the derivative $\frac{d\chi}{dx}$ of the catastrophe map in local coordinates is given by $-6x = V_a''(x)$. This leads us to the following characterization of the bifurcation set:

$$B = \{a \in \mathbb{R} \mid \exists x \in \mathbb{R} : V_a'(x) = 0, V_a''(x) = 0\}$$

This is the characterization that we actually use in the following two chapters. Since the parameter space and the state space in these chapters is multi-dimensional, the condition $V_a''(x) = 0$ has to be replaced by $\det V_a''(x) = 0$ where $V_a''(x)$ denotes the Hessian matrix at x.

Although the terms *catastrophe map* and *bifurcation set* are useful to structure our analysis of the potential function, there is a surprising result from singularity theory which helps us even more: Thom's classification theorem. It provides a local classification of all possible scenarios that can occur for potential functions with an r-dimensional parameter where $r \leq 4$. This means that we can identify the type of singularities occurring at the critical parameters which we are finding in our analysis of the bifurcation set with one of the so-called *elementary catastrophes* from Thom's list. Together with our global understanding of the bifurcation set we can then determine the precise number of global minimizers. In order to give the precise mathematical formulation of this classification theorem, we have to introduce some more notions and we follow the presentation in [45, Chapter 3].

Suppose we are given two continuous functions $f, g \colon \mathbb{R}^n \to \mathbb{R}$ both defined in a neighborhood of the origin. We say that these two functions determine the same *germ* if they coincide on some (possibly smaller) neighborhood of the origin. This defines an equivalence relation on the set of continuous functions defined on a neighborhood of the origin and

we write $\tilde{f}$ for the equivalence class of the function f. We say that a germ $\tilde{f}$ is *smooth* if it is represented by a function that is smooth in a neighborhood of the origin. For simplicity of notation, we have defined our germs to capture the infinitesimal behavior of the representatives near the origin. However, we can also speak of a germ at an arbitrary point $x \in \mathbb{R}^n$ by translating the representatives to the origin. For similar reasons, we restrict the set of germs under consideration in our exposition to the set $m(n) = \{\tilde{f} \mid f(0) = 0, f \text{ smooth}\}$. Since we are concerned with a classification of germs at local minimizers, we also introduce the set $m(n)^2 = \{\tilde{f} \in m(n) \mid f'(0) = 0\}$ of germs of functions with vanishing first derivative at the origin. In the following we will not distinguish the germ $\tilde{f}$ and the representative f of the germ anymore. Using germs instead of the functions themselves we can neglect the behavior of a function far from the local minimum under consideration. However, we also want our classification to be independent of the choice of coordinates. Therefore, we say that two germs are *equivalent* and write $f \sim g$ if there exists a smooth local diffeomorphism φ on $\mathbb{R}^n$ with $\varphi(0) = 0$ such that $f = g \circ \varphi$. We can now formulate an important property which is required for Thom's theorem:

Definition 12. *Let $f \in m(n)$ and $k \in \mathbb{N}$. We say that f is* k-determined *if for every germ $g \in m(n)$ we find $f \sim g$ whenever the k-th order Taylor polynomials of f and g near the origin coincide. If f is k-determined for some integer k, we say that f is* finitely determined*.*

Now let us come back to the example from above. The potential function $x \mapsto x^3$ clearly is a member of the family $V_a(x) = x^3 + ax, a \in \mathbb{R}$, namely for $a = 0$. Considered as a germ around the origin x^3 is a member of $m(n)$ and also $m(n)^2$. As a germ $f\colon \mathbb{R} \to \mathbb{R}$ it is also finitely-determined, in fact, it is 3-determined: Consider any other germ $g \in m(n)$ with third-order Taylor polynomial x^3 at 0. In other words, using the Fundamental Theorem of Calculus we can write $g(x) = x^3 r(x)$ in a neighborhood of 0 where

$$r(x) = \int_0^1 \int_0^1 \int_0^1 g'''(stux)\, t^2 u \, \mathrm{d}t\mathrm{d}u\mathrm{d}s.$$

If we define $\varphi(x) = x\big(r(x)\big)^{\frac{1}{3}}$, then we find $g = f \circ \varphi$, that is, $f \sim g$ if φ is a local diffeomorphism with $\varphi(0) = 0$. First note that $\varphi(0) = 0 \cdot 1^{\frac{1}{3}} = 0$

since $r(0) = 1$ and because r is continuous there is a neighborhood of 0 for which $r(x) > 0$. Furthermore, $\varphi'(0) = \big(r(0)\big)^{\frac{1}{3}} = 1 \neq 0$. Therefore we conclude that φ is a local diffeomorphism in a neighborhood of the origin. However, note that x^3 is not finitely determined as a germ $\mathbb{R}^2 \to \mathbb{R}$ since for any N such that $2N > k$ we can define a germ $g(x, y) = x^3 - y^{2N}$ that has the same k-th order Taylor polynomial as $(x, y) \mapsto x^3$ but clearly is not equivalent to it.

An *unfolding* of a germ f is a germ $F \in m(n + r)$ with some $r \in \mathbb{N}$ and $F|_{\mathbb{R}^n \times \{0\}} = f$, that is, $F(x, 0) = f(x)$ for $x \in \mathbb{R}^n$. The number r is sometimes called the *codimension* of the unfolding and we write (F, r) to state that F has codimension r. We can think of unfoldings as parameter-dependent potential functions, that is, families of potential functions. For example, in the case of the fold, the one-parameter family of potentials $(V_a)_{a\in\mathbb{R}}$ is an unfolding of the germ x^3 at $x = 0$. Similar to equivalence of germs, we can also define an equivalence relation for unfoldings. But before let us discuss an example. In addition to the unfolding $V_a(x) = x^3 + ax$ of the germ x^3 we can also consider another unfolding $G_{a,b}(x) = x^3 + ax + bx$. We can think of these unfoldings as perturbations of the germ x^3 and intuitively we might think that these two unfoldings lead to the same perturbations. We say that these two unfoldings are *associated*.

Definition 13. *Let (F, r) and (G, s) be two unfoldings of the same germ $f \in m(n)$.*

(1) F and G are associated *if there exist a smooth map Φ from $\mathbb{R}^{n+s}$ to $\mathbb{R}^{n+r}$, a smooth map ψ from $\mathbb{R}^s$ to $\mathbb{R}^r$ and a smooth real-valued function α on $\mathbb{R}^s$ such that*

   a) $\Phi(x, 0) = (x, 0)$ for every $x \in \mathbb{R}^n$,

   b) $\pi_r \circ \Phi = \psi \circ \pi_s$ where $\pi_r \colon \mathbb{R}^{n+r} \to \mathbb{R}^r$ and $\pi_s \colon \mathbb{R}^{n+s} \to \mathbb{R}^s$ are the canonical projections, and

   c) $G(x, u) = F\Big(\Phi(x, u)\Big) + \alpha(u)$.

(2) The two unfoldings F and G are said to be equivalent *if $r = s$ and the maps Φ and ψ are local diffeomorphisms.*

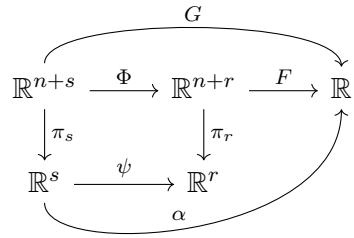

Figure 2.2: This diagram summarizes the maps involved in Definition 13. For two associated unfoldings G and F we have $G = F \circ \Phi + \alpha$.

The second part of the definition actually corresponds to the definition of *equivalent families of potential functions* in [49, Section 6.1]. The map Φ contains the parameter-dependent local diffeomorphism y_s of the state space and the map ψ plays the role of the local diffeomorphism e. Figure 2.2 summarizes the maps occurring in Definition 13.

Coming back to the example above, we can say that V and G are associated. Define the maps in the definition above as $\Phi(x,(a,b)) = (x, a+b)$, $\psi(a,b) = a+b$ and $\alpha(a,b) \equiv 0$. Then we have

$$V\Big(\Phi\big(x,(a,b)\big),\psi(a,b)\Big) + \alpha(a,b) = x^3 + (a+b)x = G(x,(a,b)).$$

So in some sense we can say that all the perturbations of G are already contained within V. This motivates the following definition. We say that an unfolding F is *stable* if any other unfolding G is associated to F. This means, intuitively, that no new bifurcation scenarios appear by adding new parameters to F.

In the second point of Definition **??**, the two unfoldings which are compared are merely reparametrizations of each other which do not change the dimension of the parameters. This will be the basis of the classification in Thom's theorem. However, we will use an even finer definition. This will be useful to recognize lower-dimensional catastrophes in higher dimensional unfoldings.

Definition 14. *Let $g \in m(n)$ and $f \in m(n+q)$. Suppose (G, r) is an unfolding of g and $(F, r+s)$ is an unfolding of f. We say that F reduces to G if $(F, r+s)$ is equivalent to the unfolding $(\hat{G}, n+q+r+s)$ with*

$$\hat{G}(x,y,u,v) = G(x,u) + y_1^2 + \cdots + y_q^2 \tag{2.27}$$

for $x \in \mathbb{R}^n, y \in \mathbb{R}^q, u \in \mathbb{R}^r$ *and* $v \in \mathbb{R}^s$.

If $q + s$ *is positive, we say that* F *reduces properly to* G *and call* G *a proper reduction of* F. *If an unfolding* F *has no proper reduction, it is called* irreducible.

If we think about the scenarios for the local minima of a family of potential functions, we might come to the conclusion that there are essentially two different ones. The first is a scenario like $x^2 + ax$ where we really have no catastrophes at all. The local minimum at $x = 0$ for $a = 0$ is merely shifted around under the variation of the parameter a. This leads to the following definition: We say that an unfolding (F, r) of $f \in m(n)$ has a *simple minimum* at 0 if f is equivalent to $x_1^2 + \cdots + x_n^2 \in m(n)$. In the other scenario, it might be that there is no local minimum at $x = 0$ at all like for the fold $x^3 + ax$. Here for $a = 0$ there is no local minimum, however if we perturb a a local minimum appears. This leads to the following definition:

Definition 15. *Let* (F, r) *be an unfolding of* $f \in m(n)$. *We say that* F *has* local minima near the origin *if for every neighborhood* W *of the origin in* $\mathbb{R}^{n+r}$ *there exists a point* $(x, u) \in W$ *such that the restriction* $F|_{(\mathbb{R}^n \times \{u\}) \cap W}$ *has a local minimum at* (x, u).

We are now prepared to state Thom's classification theorem which consists of a list of all possible scenarios for four-parameter families. For a sketch of the proof of this theorem we refer to Appendix II in [45].

Theorem 16 (Thom's classification theorem)**.** *Let* $f \in m(n)^2$ *be finitely determined,* (F, r) *a stable unfolding of* f *that has local minima near the origin and* $r \leq 4$. *Then either* F *has a simple minimum at the origin or* F *reduces to one of the irreducible unfoldings* G_i *of germs* g_i *from Table 2.1.*

We can use this theorem to get a qualitative insight into the local behavior of the potential function under investigation. Suppose that we want to understand the bifurcation set of a family of intricate potential functions V_a. If we recognize catastrophes of the Seven Elementary Catastrophes in our complicated bifurcation set, we know the local behavior of our family near those parts. This is done in Chapter 3 for the static Potts model and Chapter 4 for the time-evolved Potts model

Name	Germ g_i	Unfolding G_i	Codim.
fold	x^3	$x^3 + ux$	1
cusp	x^4	$x^4 + ux^2 + vx$	2
swallowtail	x^5	$x^5 + ux^3 + vx^2 + wx$	3
butterfly	x^6	$x^6 + ux^4 + vx^3 + wx^2 + tx$	4
hyperbolic umbilic	$x^3 + y^3$	$x^3 + y^3 + uxy + vx + wy$	3
elliptic umbilic	$x^3 - xy^2$	$x^3 - xy^2 + u(x^2 + y^2) + vx + wy$	3
parabolic umbilic	$x^2y + y^4$	$x^2y + y^4 + ux^2 + vy^2 + wx + ty$	4

Table 2.1: The Seven Elementary Catastrophes

where an incomplete (due to symmetry) butterfly unfolding and the full unfolding of the elliptic umbilic appears.

3 Stable and metastable phases for the Curie–Weiss Potts model in vector-valued fields via singularity theory

This chapter contains joint work with Prof. Dr. Christof Külske at Ruhr-Universität Bochum and is published in the Journal of Statistical Physics [38].

3.1 Introduction

3.1.1 Research context

The Potts model [59], and in particular its Curie-Weiss version, is next to the Curie-Weiss Ising model, one of the most studied models in statistical mechanics. While basic aspects of Curie-Weiss models can be discovered by ad-hoc computations, they provide ongoing challenges for refined problems involving dynamics, metastability, complex parameters, fine asymptotics [see for example 7, 8, 11, 18, 24, 41, 51]. In particular, motivated by metastability one aims at a full understanding of the free energy landscape [5, 46]. The phase-diagram for the stable states of the Curie-Weiss Potts model, that is the behaviour of *global* minimizers, is known and described by the Ellis-Wang theorem [14] in zero external field. Wang [58] provides some results in non-zero external field relating to global minimizers.

For the *metastable* states, that is, the local minima and their transitions a complete exposition of the global analysis in the whole parameter space, revealing the full structure of their transitions is lacking in the literature. See [20–22] for a partial analysis based on polynomial equivalence given

for some regions in parameter space.

One is not just interested in the static behaviour of the model, but also in the behaviour under stochastic dynamics, in and out of equilibrium, where we see next to *metastability* also the phenomenon of dynamical *Gibbs–non-Gibbs transitions* [see 30, 32, 53, 55, 57]. Employing the probabilistic notion of *sequential Gibbs property*, such dynamical Gibbs–non-Gibbs transitions have been shown to occur and analyzed for a number of exchangeable models in [31, 34, 35]. A way to understand these transitions for independent dynamics is to examine the structure of stationary points of a *conditional rate function* which generalizes the equilibrium rate function but contains more information. This conditional rate function depends on all model parameters of the static model, but has an additional time-parameter, and a measure-valued parameter with the meaning of an empirical distribution. We are planning to investigate the time-evolved Curie-Weiss Potts model in the spirit of [10, 18, 35]. For this the analysis of the static problem for *general parameter values (which means for general vector-valued fields)* is a necessary first step.

As a guiding principle for such analysis, singularity theory is very useful to understand the organization of stationary points for varying parameters. It allows to understand and discover the types of local bifurcations which are present in the applied problem as (by Thom's theorem) they must be related via local diffeomorphisms to (partial) unfoldings of elementary singularities (so-called catastrophes). These catastrophes form prototypical types which in themselves can be most easily understood in polynomial models. Clear expositions describing their geometries are found in textbooks [1, 45, 49]. The simplest example is the rate function of the Curie-Weiss Ising model. It is even globally (for all values of inverse temperature and magnetic field, and all magnetizations) identical to one cusp singularity. As we will see, for the Potts model in a general field various triple points (where three local minima have merged) occur, and singularity theory becomes very useful if we want to understand the *global picture of all the transitions the metastable minima undergo* for general values of inverse temperature and fields which allow to lift all degeneracies.

The paper is organized as follows. As our main result we describe the geometry of the bifurcation set in the parameter space of inverse temperature and vector-valued fields (see Section 3.2). This decomposition of

parameter space given by the bifurcation set provides a phase diagram describing also the metastable minima. We will show how some of the so-called elementary singularities (threefold symmetric butterflies and an elliptic umbilic surrounded by three folds) and their partial unfoldings beautifully interact and are glued together. We also describe the accompanying geometry of the free energy landscape for each connected component of the complement of the bifurcation set. Finally we complement the study with the description of the coexistence sets of lowest minima (Maxwell-sets) at which two, three, or four minima coexist (see Section 3.3).

3.1.2 Model

We consider the mean-field Potts model with three states and are interested in both the stable and metastable phases. Note that we sometimes use the term local minima to include both the global and local minima depending on the context. The space of configurations in finite-volume $n \geq 2$ is defined as $\Omega_n = \{1, 2, 3\}^n$. We define

$$\Delta^2 = \{\nu \in \mathbb{R}^3 \mid \nu_i \geq 0, \sum_{i=1}^{3} \nu_i = 1\} \tag{3.1}$$

and often refer to it as the *unit simplex*. The Hamiltonian of this model is given by

$$H_n(\sigma) = -\frac{1}{2n} \sum_{i,j=1}^{n} \delta_{\sigma_i,\sigma_j} \tag{3.2}$$

where σ lies in Ω_n. The vector-valued fields are equivalently described by the a-priori measure α in the unit simplex Δ^2. The finite-volume Gibbs distributions are therefore given by

$$\mu_{n,\beta,\alpha}(\sigma) = \frac{1}{Z_n(\beta,\alpha)} e^{-\beta H_n(\sigma)} \prod_{i=1}^{n} \alpha(\sigma_i) \tag{3.3}$$

where the partition function is defined as

$$Z_n(\beta,\alpha) = \sum_{\sigma \in \Omega_n} e^{-\beta H_n(\sigma)} \prod_{i=1}^{n} \alpha(\sigma_i). \tag{3.4}$$

So the associated free energy in terms of the empirical spin distribution ν is given by

$$f_{\beta,\alpha}(\nu) = -\frac{1}{2}\beta\langle\nu,\nu\rangle + \sum_{i=1}^{3}\nu_i \log\frac{\nu_i}{\alpha_i}. \tag{3.5}$$

This is a real-valued function on the unit simplex Δ^2 with two parameters: the inverse temperature β and the external fields modeled by the a-priori measure α. Thus, the parameter space is the product $(0,\infty)\times\Delta^2$. By a *phase* (stable or metastable) we mean a (global or local) minimizer of the free energy $f_{\beta,\alpha}$.

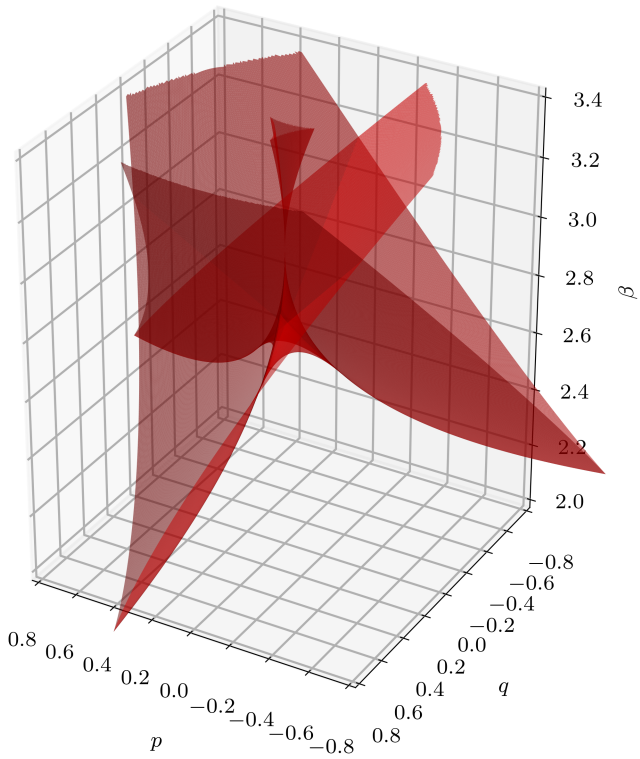

Figure 3.1: This plot shows the bifurcation set which is the basis for the metastable phase diagram. The surface shows pinches and self-intersections. Inside of the connected components of the complement the structure of metastable phases does not change. The unit simplex of a-priori measures α is embedded in the horizontal plane using the (p, q)-coordinates defined in (3.46).

3.2 The metastable phase diagram

In contrast to the usual question of phase-coexistence which is answered by the stable phase diagram (see Section 3.3), the metastable phase diagram contains information about the metastable phases of the system but not of their relative depth. Mathematically speaking, the *metastable phase diagram* is a partition of the parameter space whose cells contain parameter values (β, α) such that $f_{\beta,\alpha}$ has the same number of local minima. Using singularity theory we find that the metastable phase diagram is given by the connected complements of the surface shown in Figure 3.1. The structure of this union is particularly interesting because it shows features of two well-known catastrophes [49]: the butterfly catastrophe and the elliptic umbilic. The elliptic umbilic permits the change from minimum to maximum at the centre and is inherently associated to the Potts model because of its symmetry. The appearance of the butterfly is connected to triple points of Ising-like subsystems of the three-state Potts model. If we disfavor one of the three states, the remaining two act similarly to an Ising model in a random field. There is an interesting global interdependence of these two different catastrophes. We summarize the geometry of the extended phase diagram in the following theorem.

Theorem 17. *For each positive β define the so-called catastrophe map χ_β from Δ^2 to Δ^2 which associates to each empirical spin distribution a β-dependent a-priori measure α modelling the external fields:*

$$\chi_\beta(\nu) = \left(\frac{\nu_i e^{-\beta\nu_i}}{\sum_{k=1}^{3} \nu_k e^{-\beta\nu_k}} \right)_{i=1}^{3} \tag{3.6}$$

Define also curves γ_β via

$$\gamma_\beta(x) = \frac{1}{2}\left(1 - x - \sqrt{(1-x)^2 - \frac{4(1-2\beta x(1-x))}{\beta(2-3\beta x)}}\right) \tag{3.7}$$

for $x \in D_\beta$ where the domain D_β is a union of intervals given by

$$D_\beta = \begin{cases} \left(0, 1-\frac{2}{\beta}\right] \cup \left(\frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{2}{\beta}}, \frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{2}{\beta}}\right) & \text{if } 2 < \beta \leq \frac{8}{3}, \\ \left(0, \frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{2}{\beta}}\right) \cup \left(1-\frac{2}{\beta}, \frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{8}{3\beta}}\right) & \\ \cup \left(\frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{8}{3\beta}}, \frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{2}{\beta}}\right) & \text{if } \frac{8}{3} \leq \beta < 3, \\ \left(0, \frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{2}{\beta}}\right) \cup \left(\frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{8}{3\beta}}, 1-\frac{2}{\beta},\right) & \\ \cup \left(\frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{8}{3\beta}}, \frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{2}{\beta}}\right) & \text{if } 3 \leq \beta. \end{cases} \tag{3.8}$$

Then consider the curve Γ_β in Δ^2 given by $\Gamma_\beta(x) = \chi_\beta(x, \gamma_\beta(x), 1-x-\gamma_\beta(x))$ with $x \in D_\beta$. By $S_3\Gamma_\beta(D_\beta)$ we denote the orbit of the curve Γ_β under the action of the permutation group S_3 acting on Δ^2.

(1) The constant-temperature slices of the bifurcation set from Figure 3.1 are given by $S_3\Gamma_\beta(D_\beta)$.

(2) Table 3.1 shows an overview of the number of connected components of $S_3\Gamma_\beta(D_\beta)$ together with the possible number of local minima. The exact number of local minima in the respective components can be seen in Figure 3.2.

3.2.1 Main transitions

Increasing the inverse temperature from zero we see the following transitions for slices of the bifurcation set at fixed inverse temperature, as plots of curves in two-dimensional magnetic field space (see Figure 3.2). In connected complements the structure of stationary points does not change. We will keep track of the number of minimizers, which takes all values between one and four. The Maxwell sets where non-uniqueness of global minimizers occurs are described in Section 3.3. They have the meaning of *special magnetic fields* where (in general) the minimizer can be made to jump by infinitesimal perturbation. This is analogous to the notion of *bad empirical measures* in the dynamical model, in the sense of sequential Gibbsianness [35].

- $2 < \beta < \frac{18}{7}$. First *three symmetric cusps* appear at a positive distance to the origin in magnetic field space (three "rockets" pointing

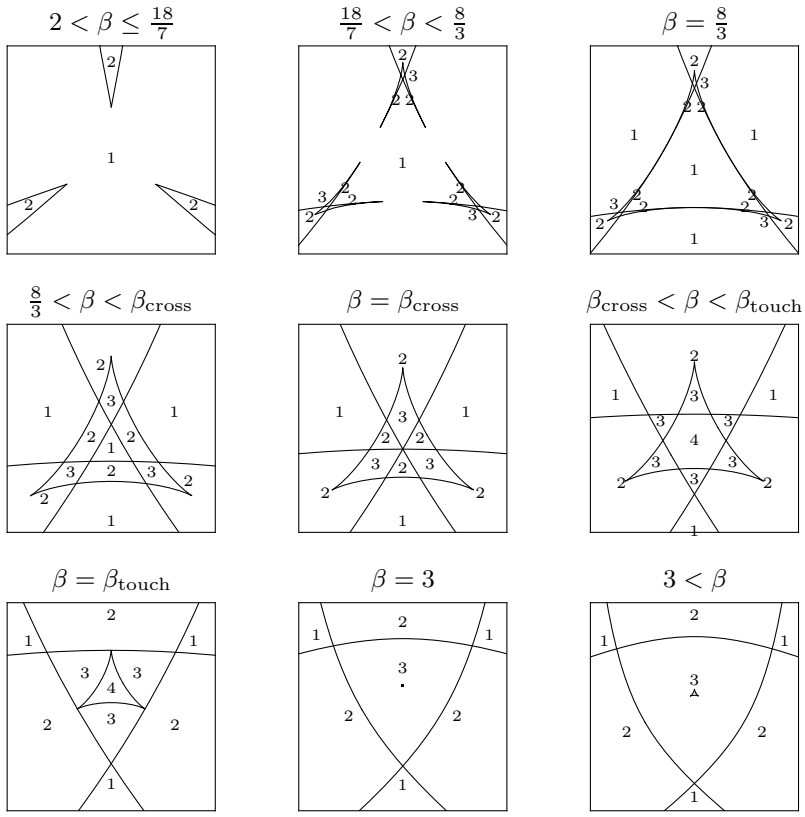

Figure 3.2: Representative slices through the bifurcation set at constant temperatures as indicated by the plot titles. The numbers of the cells show the number of local minima that the rate function $f_{\beta,\alpha}$ has inside of the respective cell. The slices are given in (p, q)-coordinates (3.46) as in Figure 3.1.

	cells of $(S_3\Gamma_\beta(D_\beta))^{\complement}$	number of local minima
$\beta \leq 2$	1	1
$2 < \beta \leq \frac{18}{7}$	4	1, 2
$\frac{18}{7} < \beta < \frac{8}{3}$	13	1, 2, 3
$\beta = \frac{8}{3}$	16	1, 2, 3
$\frac{8}{3} < \beta < \beta_{\text{cross}}$	13	1, 2, 3
$\beta = \beta_{\text{cross}}$	12	1, 2, 3
$\beta_{\text{cross}} < \beta < \beta_{\text{touch}}$	13	1, 2, 3, 4
$\beta = \beta_{\text{touch}}$	10	1, 2, 3, 4
$\beta_{\text{touch}} < \beta < 3$	8	1, 2, 3
$\beta = 3$	7	1, 2, 3
$3 < \beta$	8	1, 2, 3

Table 3.1: Overview of the different β-regimes for the constant-temperature slices of the bifurcation set. The critical inverse temperatures β_{cross} and β_{touch} are defined in Subsections 3.2.4 and 3.2.4.

towards the origin). For magnetic fields inside the cusps we see precisely two minima, outside there is one minimum. For each such inverse temperature, the effect of the two-dimensional magnetic field for values in the interior of this region to this effective orthogonal Ising model translates into an effective inverse temperature times effective magnetic field.

- $\beta = \frac{18}{7}$ (butterfly). The three cusps each individually develop a butterfly singularity. The unfolding of the pentagram-shaped curve is studied via a Taylor expansion in Subsection 3.2.4.

- $\frac{18}{7} < \beta < \frac{8}{3}$. The butterfly (partially) unfolds, keeping the reflection symmetry. This phenomenon is also known to occur in the Curie-Weiss random field Ising model with bimodal disorder [compare 35]. The potential has two minima in the outer horns of the pentagram, and three minima in the inner horn, as known from the one-dimensional polynomial model. For zero magnetic field there is still one minimum in the centre of the simplex.

- $\beta = \frac{8}{3}$. The outer horns (or beaks) of the pentagrams grow until

they meet symmetrically in a *beak-to-beak* singularity. This occurs in three pairs. A one-pair beak-to-beak singularity is also known to occur in the parabolic umbilic [see 6]. This touching creates a finite connected component at the origin in magnetic field space, still with one minimum.

- $\frac{8}{3} < \beta < \beta_{\text{cross}} \approx 2.7456$. Each two of the beaks (outer horns corresponding to different butterflies) have now become connected. Each such pair now forms a joint connected component with two minima. Their three outer boundary curves now form a triangle in the centre. Each two of the former unbounded curves of the butterfly catastrophe have merged into one doubly infinite curve at which a fold occurs (fold lines). In the connected component at the origin there is still one minimum.

- $\beta = \beta_{\text{cross}}$. The triangle stays, the three symmetric fold lines move towards the origin. They pass the origin at β_{cross}, when the "tops of the rockets" meet at the origin, and the connected component containing the origin with one minimum vanishes. β_{cross} is the parameter value for the appearance of symmetric minima near the corners in zero magnetic field. Hence it is simply found by looking at the potential in zero field, along the axis of symmetry (see 3.2.4).

- $\beta_{\text{cross}} < \beta < \beta_{\text{touch}} \approx 2.8024$. The three rockets move on beyond the origin, they intersect, with the *appearance of a middle hexagon*. In this middle hexagon containing the origin there are all four minima present. In zero field the middle minimum is the lowest first but moving beyond the Ellis-Wang critical inverse temperature $4 \log 2$, eventually the outer minima become lower. In the adjacent six triangles there are three minima.

- $\beta \geq \beta_{\text{touch}}$. Three components with three minima vanish, three components remain, as the corners of the shrinking triangle touch the fold lines.

- $\beta = 3$ (elliptic umbilic). The triangle at the centre has shrunk to a point, the minimum at zero in zero field has become a monkey saddle.

- $\beta > 3$. The inner triangle reappears and grows again. For zero field there is a maximum at the uniform distribution, and three symmetric minimizers near the corners.

The series of transitions upon increasing inverse temperature fits to the basic knowledge of the model without fields [14]: We know that in zero field a) at very low inverse temperature there is only one local minimum (and this is also a global minimum) at the uniform distribution, b) at intermediate inverse temperature there is a local minimum at the uniform distribution and three symmetric minima near the corners c) at large β there are only three symmetric minima near the corners.

The change from minimum to maximum of the uniform distribution under increase of β is explained by an elliptic umbilic. Additionally, for each of the minima at the corners there is an additional fold line. There must be a transition from the situation of three non-intersecting rockets $\beta \approx 2.2$ to an umbilic plus three fold lines seen at the Ellis-Wang inverse temperature $4 \log 2$. This is done with the help of the three-symmetric-butterflies – beak-to-beak mechanism.

3.2.2 Elements from singularity theory

In order to derive and explain our results, concepts from singularity theory will be useful. The two most basic terms are *catastrophe manifold* and *bifurcation set* of which the second term is important since it is the basis for the metastable phase diagram. But first let us define the two: The *catastrophe manifold* is the set of (β, α, ν) such that ν is a stationary point of $f_{\beta,\alpha}$. The *bifurcation set* is the set of (β, α) such that there exists a degenerate stationary point ν for $f_{\beta,\alpha}$, that is, a stationary point at which the Hessian has a zero eigenvalue. The *catastrophe map* χ_β maps empirical spin distributions ν to a-priori measures α such that the free energy $f_{\beta,\alpha}$ has a stationary point at ν. We obtain the expression for the catastrophe map by considering the zeros of the differential of $f_{\beta,\alpha}$. For every tangent vector v of the unit simplex we have

$$\sum_{i=1}^{3} v(\nu_i)(-\beta\nu_i - \log\alpha_i + \log\nu_i) = 0. \tag{3.9}$$

We conclude that the second factor in the sum is a constant since the $v(\nu_i)$ sum up to zero. Since α is an element of the unit simplex, we have Equation (3.6).

The key idea from catastrophe theory is that at parameter values belonging to the bifurcation set the stationary points of the function change. The most generic change is the fold where a minimum and a maximum collide. But the more parameters the potential has the more complex behaviour is possible. The famous theorem of Thom [see Section 5 of Chapter 3 in 45] lists these possibilities for all potentials with at most four parameters. We see two of these so-called *catastrophes* or *singularities* in the Potts model: the butterfly catastrophe and the elliptic umbilic.

3.2.3 Constant-temperature slices of the bifurcation set

The constant-temperature slices of the bifurcation set are one-dimensional sets in the sense that we have a parametric representation of the slices with one parameter but the curves show pinches and self-intersections. During the computation of these curves we can already see some critical behaviour corresponding to the beak-to-beak scenario and the elliptic umbilic point.

It is convenient to study the degenerate stationary points for fixed temperature first. From these points we can obtain the respective slice of the bifurcation set via the catastrophe map χ_β.

Theorem 18. *The set of degenerate stationary points for any positive β is given by the symmetrized graph of the smooth function γ_β on D_β defined in Equations (3.7) and (3.8).*

By the symmetrized graph of γ_β we mean the orbit $S_3\gamma_\beta(D_\beta)$ under the permutation group S_3. Observe the critical behaviour that $D_\beta = (0,1)$ for $\beta = \frac{8}{3}$ (beak-to-beak) and that the two roots $1-\frac{2}{\beta}$ and $\frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{8}{3\beta}}$ coincide for $\beta = 3$ (Elliptic umbilic). We will now provide a series of lemmata in preparation of the proof of this theorem.

The set of degenerate stationary points is determined by the so-called *degeneracy condition*. This condition states that the determinant of the Hessian matrix at a stationary point vanishes.

Lemma 19. *The Hessian form of $f_{\beta,\alpha}$ at the stationary point ν is degenerate if and only if*

$$3\nu_1\nu_2\nu_3\beta^2 - 2(\nu_1\nu_2 + \nu_2\nu_3 + \nu_3\nu_1)\beta + 1 = 0. \tag{3.10}$$

Proof. Let ν be a stationary point. Choose (ν_1, ν_2) as local coordinates for ν. The Hessian form at ν is represented with respect to the coordinate basis by the matrix

$$\begin{pmatrix} \frac{1}{\nu_1} + \frac{1}{\nu_3} - 2\beta & \frac{1}{\nu_3} - \beta \\ \frac{1}{\nu_3} - \beta & \frac{1}{\nu_2} + \frac{1}{\nu_3} - 2\beta \end{pmatrix}.$$

Calculating the determinant yields

$$3\beta^2 - 2\beta\left(\frac{1}{\nu_1} + \frac{1}{\nu_2} + \frac{1}{\nu_3}\right) + \frac{1}{\nu_1\nu_2} + \frac{1}{\nu_1\nu_3} + \frac{1}{\nu_2\nu_3}.$$

□

If we rewrite the left-hand side of the degeneracy condition (3.10) in (ν_1, ν_2)-coordinates, it is a quadratic function of ν_2 for fixed β and ν_1:

$$\beta(2 - 3\beta\nu_1)\nu_2^2 - \beta(2 - 3\beta\nu_1)(1 - \nu_1)\nu_2 + 1 - 2\beta\nu_1(1 - \nu_1) = 0 \tag{3.11}$$

This equation has at most two solutions and one of them is given by

$$\nu_2 = \gamma_\beta(\nu_1). \tag{3.12}$$

The possible other solution is obtained by applying the respective symmetry operation (exchanging the second and third component of ν). The domain of γ_β is determined by the sign of the discriminant and the additional condition that makes sure that the result of γ_β is a point in the unit simplex. Let us investigate the latter condition first (Lemma 20) and come to the condition imposed by nonnegativity of the discriminant (Lemma 21) afterwards.

First note that the solution formula (3.7) is not defined for $x = \frac{2}{3\beta}$ but converges to $\pm\infty$ because

$$1 - 2\beta x(1 - x) = \frac{1}{3}\left(\frac{8}{3\beta} - 1\right)\begin{cases} > 0 & \beta < \frac{8}{3} \\ = 0 & \beta = \frac{8}{3} \\ < 0 & \beta > \frac{8}{3} \end{cases} \tag{3.13}$$

except if $\beta = \frac{8}{3}$ where the limit is $\frac{1}{4}$. Furthermore, the domain of γ_β must be such that $(x, \gamma_\beta(x), 1 - x - \gamma_\beta(x))$ lies in the unit simplex, that is, we have to analyse the following system of inequalities:

$$\begin{aligned} 0 < x < 1 \\ 0 < \gamma_\beta(x) < 1 - x \end{aligned} \tag{3.14}$$

Lemma 20. *Provided that $\gamma_\beta(x)$ is a real number, we find:*

(1) For $\beta < \frac{8}{3}$ the system (3.14) is equivalent to

$$x \in \left(0, \frac{2}{3\beta}\right) \cup \left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2}{\beta}}, \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2}{\beta}}\right) \tag{3.15}$$

(2) For $\frac{8}{3} \leq \beta$ the system (3.14) is equivalent to

$$x \in \left(0, \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2}{\beta}}\right) \cup \left(\frac{2}{3\beta}, \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2}{\beta}}\right) \tag{3.16}$$

Proof.

$$0 < \gamma_\beta(x) < 1 - x \tag{3.17}$$

$$\iff -(1-x) < \sqrt{(1-x)^2 - \frac{4(1 - 2\beta x(1-x))}{\beta(2 - 3\beta x)}} < 1 - x \tag{3.18}$$

The first inequality is trivially true because the square root is non-negative and $1 - x > 0$. Therefore we must only check the second inequality:

$$\sqrt{(1-x)^2 - \frac{4(1 - 2\beta x(1-x))}{\beta(2 - 3\beta x)}} < 1 - x \tag{3.19}$$

$$\iff 0 < \frac{1 - 2\beta x(1-x)}{2 - 3\beta x} \tag{3.20}$$

$$\tag{3.21}$$

Suppose first $2 - 3\beta x > 0$. Then the inequality is equivalent to

$$1 - 2\beta x(1-x) > 0 \tag{3.22}$$

$$\iff 2\beta\left(x - \frac{1}{2}\right)^2 + 1 - \frac{\beta}{2} > 0 \tag{3.23}$$

As $2\beta > 0$, this is an upfacing parabola whose minimal functional value is $1 - \frac{\beta}{2}$ which is negative since $\beta > 2$. The roots of this parabola are:

$$2\beta(x - \frac{1}{2})^2 + 1 - \frac{\beta}{2} = 0 \tag{3.24}$$

$$\iff x = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{2}{\beta}} \tag{3.25}$$

The solution is therefore the union of the two intervals $0 < x < \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2}{\beta}}$ and $\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2}{\beta}} < x < 1$.

Now, if $2 - 3\beta x < 0$, we are looking for such x that the values of $1 - 2\beta x(1-x)$ are negative. This is the case if

$$\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2}{\beta}} < x < \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2}{\beta}}. \tag{3.26}$$

To arrive at the claim in the lemma, we must investigate the order of $\frac{2}{3\beta}$ and $\frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{2}{\beta}}$. We therefore analyse the inequality

$$\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2}{\beta}} > \frac{2}{3\beta}. \tag{3.27}$$

Squaring both sides of the inequality reveals that it is equivalent to $\beta < \frac{8}{3}$ which proves the claim. □

Let us continue with the analysis of the discriminant of the quadratic equation (3.11). It is given by

$$\begin{aligned}\beta^2(2 - 3\beta\nu_1)^2(1-\nu_1)^2 - 4\beta(2 - 3\beta\nu_1)(1 - 2\beta\nu_1(1-\nu_1)) \\ = \beta(3\beta\nu_1 - 2)(2 - \beta(1-\nu_1))(2 - 3\beta\nu_1(1-\nu_1)).\end{aligned} \tag{3.28}$$

We see that it has at most four possible roots depending on the value of β. More precisely we find:

Lemma 21. *Consider the function*

$$x \mapsto (3\beta x - 2)(2 - \beta(1-x))(2 - 3\beta x(1-x)) \tag{3.29}$$

for real x.

(1) For $\beta < \frac{8}{3}$ this function has the two roots $1 - \frac{2}{\beta} < \frac{2}{3\beta}$ and takes positive values only on $(-\infty, 1 - \frac{2}{\beta}) \cup (\frac{2}{3\beta}, \infty)$.

(2) For $\frac{8}{3} \leq \beta < 3$ this function has the roots $\frac{2}{3\beta} \leq 1 - \frac{2}{\beta} < \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}} \leq \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}$ and takes positive values only on

$$\left(-\infty, \frac{2}{3\beta}\right) \cup \left(1 - \frac{2}{\beta}, \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}\right) \cup \left(\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}, \infty\right).$$

The equality of the roots is achieved for $\beta = \frac{8}{3}$.

(3) For $\beta \geq 3$ this function has the roots $\frac{2}{3\beta} < \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}} \leq 1 - \frac{2}{\beta} < \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}$ and takes positive values only on

$$\left(-\infty, \frac{2}{3\beta}\right) \cup \left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}, 1 - \frac{2}{\beta}\right) \cup \left(\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}, \infty\right)$$

The equality of the roots is achieved for $\beta = 3$.

Proof. The expression for the roots follow from the product form of the function. Since we know all roots, the set where the function takes positive values is determined by the sign change at the roots. First, let us consider $\beta < \frac{8}{3}$. This implies the order of the roots since $1 - \frac{2}{\beta} < \frac{1}{4} < \frac{2}{3\beta}$. The value of the derivative at the roots tells us how the sign changes. Denote the above function (3.29) by g. Then we find:

$$g'\left(\frac{2}{3\beta}\right) = 4\left(\frac{8}{3} - \beta\right)$$
$$g'\left(1 - \frac{2}{\beta}\right) = 12\left(\beta - \frac{8}{3}\right)(3 - \beta)$$

This proves the case $\beta < \frac{8}{3}$.

Secondly, consider the case $\frac{8}{3} \leq \beta < 3$ but first assume $\beta > \frac{8}{3}$. The order of the of the roots $\frac{2}{3\beta}$ and $1 - \frac{2}{\beta}$ now reverses and $\beta < 3$ implies $1 - \frac{2}{\beta} < \frac{1}{3} < \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}$. Let us now analyse how the sign changes at the two largest roots. Let x_0 be any of the two roots of $2 - 3\beta x(1-x)$. We find

$$g'(x_0) = (3\beta x_0 - 2)(2 - \beta(1 - x_0)) \cdot 3\beta(2x_0 - 1). \tag{3.30}$$

Since $3\beta x - 2$ and $2 - \beta(1-x)$ are increasing and x_0 is larger than their roots, the sign of $g'(x_0)$ is determined by the sign of $2x_0 - 1$ which is negative for $x_0 = \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}$ and positive for $x_0 = \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}$. In the case $\beta = \frac{8}{3}$ the roots $\frac{2}{3\beta}$ and $1 - \frac{2}{\beta}$ as well as the two largest roots coincide and for both the sign does not change.

Let us now consider the last case $3 \leq \beta$ focusing on $3 < \beta$ first. Let us check the order of the roots. Using the inequality $\sqrt{1+x} < 1 + \frac{1}{2}x$ for $x < 0$, we find that $\frac{2}{3\beta} < \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}$. Similarly to the previous case the reversed inequality $\beta > 3$ now implies the reversed inequality $\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}} < \frac{1}{3} < 1 - \frac{2}{\beta}$. The last inequality for the roots is in fact equivalent to $3 < \beta$:

$$1 - \frac{2}{\beta} < \frac{1}{2}\left(1 + \sqrt{1 - \frac{8}{3\beta}}\right) \tag{3.31}$$

$$\iff 1 + \frac{16}{\beta^2} - \frac{8}{\beta} < 1 - \frac{8}{3\beta} \tag{3.32}$$

$$\iff 3 < \beta \tag{3.33}$$

Let us analyse how the sign changes at the two roots of $2 - 3\beta x(1-x)$. This can be done using formula (3.30). The first factor is positive for both roots since $\frac{2}{3\beta}$ is the smallest root of the discriminant. However, since $1 - \frac{2}{\beta}$ lies in between the two roots of $2 - 3\beta x(1-x)$, the sign of g' is positive at both roots $\frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}$. For $\beta = 3$ the roots $1 - \frac{2}{\beta}$ and $\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{8}{3\beta}}$ coincide and form a local maximum. □

Proof of Theorem 18. The expression for $\gamma_\beta(x)$ (formula (3.7)) is obtained by solving the quadratic equation (3.11) for ν_2. In order to

determine the domain D_β we have to intersect the sets from Lemma 20, which ensures that $(x, \gamma_\beta(x), 1 - x - \gamma_\beta(x))$ lies in the unit simplex, and Lemma 21 ensuring the nonnegativity of the discriminant. □

3.2.4 Computation of the critical temperatures

We discuss the various critical temperatures in increasing order.

The butterfly temperature

Looking at the constant-temperature slices of the bifurcation set in the regime $2 < \beta < \frac{8}{3}$, we find a qualitative change of the curve (compare Figure 3.2): A pentagram-like shape unfolds. The butterfly temperature is defined by the β at which this happens which is for $\beta = \frac{18}{7}$. This can be seen by a Taylor expansion of the curve which describes the constant-temperature slices as the coefficients undergo sign changes. Because of symmetry, it does not matter which of the three rockets we consider. Let us consider the degenerate stationary points with $\nu_3 \leq \min\{\nu_1, \nu_2\}$. More precisely, consider the degeneracy equation (3.10) in the following coordinates:

$$\begin{aligned} x &= \frac{\sqrt{3}}{2}(\nu_1 - \nu_2) \\ y &= \frac{1}{2}(3\nu_3 - 1) \end{aligned} \tag{3.34}$$

In these coordinates the unit simplex Δ^2 is an equilateral triangle with center at the origin. The equation then reads

$$\frac{1}{9}\left(6\beta\left(x^2 + y^2 - 1\right) + \beta^2(2y+1)\left((y-1)^2 - 3x^2\right) + 9\right) = 0. \tag{3.35}$$

For $x = 0$, this equation reads

$$2\beta^2 y^3 + 3\beta(2-\beta)y^2 + (\beta - 3)^2 = 0. \tag{3.36}$$

which has a single, negative root y_0. Using the Implicit Function Theorem the set of degenerate stationary points locally around $(0, y_0)$ is the graph of a function $y = g_\beta(x)$ which solves (3.35). It is of course also possible to obtain the values of $g_\beta^{(n)}(0)$ using Implicit Differentiation which allows

us to write down a Taylor expansion for g_β. If we plug this into the catastrophe map χ_β we arrive at an expansion for the respective slice of the bifurcation set:

$$\begin{aligned}
&\chi_\beta(x, g_\beta(x)) - \chi_\beta(0, y_0) = \\
&-\binom{1}{1}\left(\frac{27}{2}\left(\beta - \frac{18}{7}\right) + O\left(\left(\beta - \frac{18}{7}\right)^2\right)\right)x^2 + \\
&\quad\binom{-1}{1}\left(\frac{36}{7}\sqrt{3}\left(\beta - \frac{18}{7}\right) + O\left(\left(\beta - \frac{18}{7}\right)^2\right)\right)x^3 + \\
&\quad\binom{1}{1}\left(\frac{39366}{2401} - \frac{48519}{343}\left(\beta - \frac{18}{7}\right) + O\left(\left(\beta - \frac{18}{7}\right)^2\right)\right)x^4 + O\left(x^5\right)
\end{aligned} \tag{3.37}$$

This has been achieved using exact computations in Mathematica. Here we have slightly abused notation by writing χ_β for the coordinate representation of χ_β using (x, y)-coordinates in the source and the following coordinates in the target space:

$$\begin{aligned}
u &= \log\frac{\alpha_1}{\alpha_3} \\
v &= \log\frac{\alpha_2}{\alpha_3}
\end{aligned} \tag{3.38}$$

So the coordinate representation of the catastrophe map χ_β is given by

$$(x, y) \mapsto \begin{pmatrix} \log\frac{1-y-\sqrt{3}x}{1+2y} + \beta\left(y + \frac{x}{\sqrt{3}}\right) \\ \log\frac{1-y+\sqrt{3}x}{1+2y} + \beta\left(y - \frac{x}{\sqrt{3}}\right) \end{pmatrix}. \tag{3.39}$$

The crossing temperature

Lemma 22. *The inverse crossing temperature is given by*

$$\beta_{\text{cross}} = \frac{3}{(1+2s)(1-s)} \approx 2.74564 \tag{3.40}$$

where s is the unique root in $(0, 1)$ of

$$\frac{3x}{(1+2x)(1-x)} - \log\frac{1+2x}{1-x}. \tag{3.41}$$

Proof. Let $\gamma(s) = \frac{1}{3}(1+2s, 1-s, 1-s)$ for $0 \leq s < 1$ and let α be the uniform distribution. The inverse crossing temperature β_{cross} equals the β such that the two outer local extrema of $f_{\beta,\alpha} \circ \gamma$ annihilate. This is characterized by the two equations (first and second derivatives of $f_{\beta,\alpha} \circ \gamma$)

$$\beta s - \log \frac{1+2s}{1-s} = 0 \tag{3.42}$$

$$\beta - \frac{3}{(1+2s)(1-s)} = 0. \tag{3.43}$$

Plugging (3.43) into (3.42) motivates the following definition: Let g be given on $(0,1)$ by

$$g(x) = \frac{3x}{(1+2x)(1-x)} - \log \frac{1+2x}{1-x}. \tag{3.44}$$

The first derivative of this function vanishes in the open interval $(0,1)$ exactly at $x = \frac{1}{4}$, it is decreasing on $\left(0, \frac{1}{4}\right)$, increasing on $\left(\frac{1}{4}, 1\right)$ and $\lim_{x\to 0} g(x) = 0$. Therefore g has a unique root in $(0,1)$. □

The triangle-touch temperature

The *triangle-touch temperature* is defined as the temperature $1/\beta_{\text{touch}}$ such that in the respective constant-temperature slice the vertices of the central triangle touch the fold lines. By definition, $\frac{8}{3} < \beta_{\text{touch}} < 3$.

Lemma 23. *The inverse triangle-touch temperature β_{touch} is the unique zero in $[\frac{8}{3}, 3]$ of*

$$2\,\operatorname{artanh}\sqrt{1-\frac{8}{3x}} + \frac{3x}{4}\left(1-\sqrt{1-\frac{8}{3x}}\right) - \log\left(\frac{x-2}{2}\right) - 3. \tag{3.45}$$

Proof. First, observe that the function is strictly increasing, positive for $x = \frac{8}{3}$ and negative for $x = 3$. The function values are $1 - \log 3$ and $\frac{3}{2} - \log 4$ respectively. Therefore this function has a unique zero in the specified interval.

It suffices to show that one vertex of the central triangle and one of the fold lines meet because of symmetry. Since the vertex lies on an

axis of symmetry of the simplex for all $\frac{8}{3} < \beta < 3$, we know that the intersection point with the fold line must also lie on the axis of symmetry (the centre of the fold line). For the space of a-priori measures α we use the following coordinates:

$$\begin{aligned} p &= \sqrt{3} \log \frac{\alpha_1}{\alpha_2} \\ q &= \log \frac{\alpha_1 \alpha_2}{\alpha_3^2} \end{aligned} \tag{3.46}$$

The vertex of the triangle fulfills $p = \sqrt{3}(\log(\beta - 2) + 3 - \beta)$ and the centre of the fold line has

$$p = \sqrt{3} \left(\log 2 - \frac{1}{4}\beta - \frac{3\beta}{4} \sqrt{1 - \frac{8}{3\beta}} + 2 \operatorname{artanh} \sqrt{1 - \frac{8}{3\beta}} \right).$$

Equating the two formulas proves the claim. The values of the u-coordinate can be calculated from the respective degenerate stationary points (see Figure 3.3). The ν_1 values are the lower bounds of the domains and $\nu_2 = \frac{1}{2}(1 - \nu_1)$. □

The elliptic umbilic temperature

We know from singularity theory that the elliptic umbilic is a doubly degenerate point, that is, a point where the Hessian has two zero eigenvalues. There is only one such point for this potential and it is given by $\beta = 3$: The vanishing mixed second-order partial derivatives of the potential implies $\beta = \frac{1}{\nu_3}$. Plugging this into

$$\frac{\partial^2 f_{\beta,\alpha}}{\partial \nu_i^2} = 2\beta - \frac{1}{\nu_i} - \frac{1}{\nu_3} = 0$$

yields $\frac{1}{\nu_i} = \beta$. These partial derivatives vanish therefore only for $\beta = 3$ and $\nu = \frac{1}{3}(1, 1, 1)$ in zero magnetic field. Furthermore, the third-order Taylor expansion of the potential $f_{\beta,\alpha}$ for $\beta = 3$ and zero magnetic field in the (x, y)-coordinates (3.34) is given by $-\frac{1}{3}y^3 + x^2 y + \text{const}$. This is exactly the germ of the elliptic umbilic from Thom's seven elementary catastrophes.

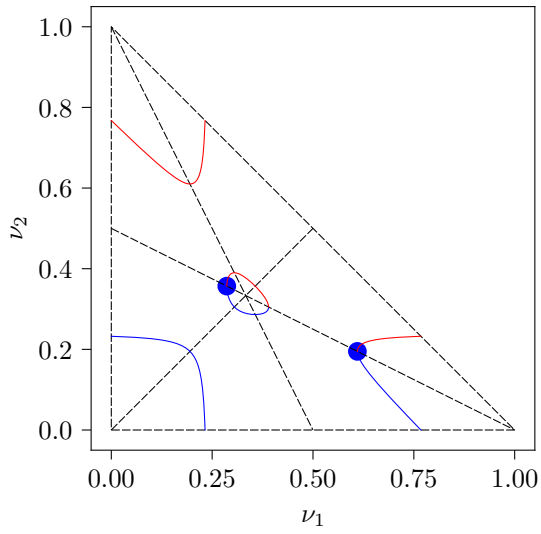

Figure 3.3: The figure shows the two degenerate stationary points in ν-space that are mapped to a touching point of the inner triangle with of a fold line in α-space under the catastrophe map χ_β.

3.2.5 A parametric representation of the bifurcation set

As we have learned in the previous subsections, the extended phase diagram is constructed via the bifurcation set. In this subsection we present the parametric representation which was used to create Figure 3.1 and how it is obtained.

Theorem 24. *The bifurcation set is given by the union $F^+(\Delta^2) \cup F^-(\Delta^2)$ of the images of the two maps $F^\pm$ from Δ^2 to the parameter space $(0,\infty) \times \Delta^2$ with components*

$$F_1^\pm(\nu) = \frac{1}{3}\sum_{i=1}^{3}\frac{1}{\nu_i} \pm \sqrt{\frac{1}{9}\left(\sum_{i=1}^{3}\frac{1}{\nu_i}\right)^2 - \frac{1}{3}\sum_{i<j}\frac{1}{\nu_i\nu_j}} \tag{3.47}$$

$$F_2^\pm(\nu) = \chi_{F_1^\pm(\nu)}(\nu) = \left(\frac{\nu_i e^{-F_1^\pm(\nu)\nu_i}}{\sum_{k=1}^{q}\nu_k e^{-F_1^\pm(\nu)\nu_k}}\right)_{i=1}^{3} \tag{3.48}$$

Proof. First, let us check that F^+ and F^- map Δ^2 into $(0,\infty) \times \Delta^2$.

Clearly, $F_2^{\pm}(\nu)$ is an element of Δ^2. Furthermore, $F_1^{+}(\nu)$ is obviously positive. Since every ν_i is positive, we have

$$\sqrt{\frac{1}{9}\left(\sum_{i=1}^{3}\frac{1}{\nu_i}\right)^2 - \frac{1}{3}\sum_{i<j}\frac{1}{\nu_i\nu_j}} < \frac{1}{3}\sum_{i=1}^{3}\frac{1}{\nu_i} \tag{3.49}$$

which implies that $F_1^{-}(\nu)$ is also positive. A point (β, α) belongs to the bifurcation set if and only there exists a degenerate stationary point ν in the unit simplex such that $\alpha = \chi_\beta(\nu)$. Lemma 19 shows us that the degeneracy condition (3.10) is a quadratic equation but this time considered as a function of β. Note also that it is independent of α. The discriminant is given by

$$4(\nu_1\nu_2 + \nu_2\nu_3 + \nu_3\nu_1)^2 - 12\nu_1\nu_2\nu_3 \tag{3.50}$$

$$= 4((\nu_1\nu_2 + \nu_2\nu_3 + \nu_3\nu_1)^2 - 3\nu_1\nu_2\nu_3) \tag{3.51}$$

$$= 4((\nu_1\nu_2)^2 + (\nu_2\nu_3)^2 - \nu_1\nu_2\nu_3 + (\nu_3\nu_1)^2). \tag{3.52}$$

On the boundary of the simplex, the discriminant is positive except at the vertices. We see by calculus that it achieves its minimal value at the centre of the simplex where it takes the value zero. Therefore the quadratic equation for β has the two solutions $F_1^{\pm}(\nu)$ for every ν.

Now that we know the two possible β-values that make ν fulfill the degeneracy condition we can use the catastrophe map χ_β to obtain the respective a-priori measures α to make them degenerate stationary points. This proves the claim. □

3.3 The stable phase diagram

The classical phase diagram is a partition of the parameter space. However, in contrast to the metastable phase diagram, the cells of this partition contain (β, α) such that the number of *global* minimizers stays constant inside the cell. This type of phase diagram does not see the fine bifurcation behaviour of the rate function and is therefore much simpler to describe. We can think of the classical phase diagram as given by a "surface" in the parameter space: the coexistence surface. On this surface we have a coexistence of at least two phases. It is therefore clear

that the coexistence surface lies in those cells of the metastable phase diagram in which the rate function has at least two minimizers. The complement of the coexistence surface defines a region of the parameter space in which the rate function has a unique global minimum.

The surface is best understood by moving in the direction of increasing β (see Figures 3.4 to 3.7). Leaving the high temperature regime ($\beta \leq 2$), it consists of three lines on the axis of symmetry. These lines have progeny, namely two lines emerging at a positive angle. This results in three Y-shaped sets composed of curved and straight lines. Furthermore, each of the offspring lines of the Y-shaped curves meet during the beak-to-beak scenario we have already seen in the extended phase diagram and form a triangle. Finally, the triangle shrinks to a point and we see a star-shaped set consisting of three straight lines (the axes of symmetry). On the coexistence surface the rate function has at least two and at most four global minimizers. The point of coexistence of four phases is the well-known Ellis-Wang point [14]. For $2 < \beta \leq \frac{18}{7}$ it is only possible to have a coexistence of two phases. However, starting with β larger than $\frac{18}{7}$ we find so-called triple points (coexistence of three phases). These will be important for our numerical computation of the Maxwell set (coexistence surface). Let us summarize our result:

Theorem 25. *Let us use the* (x, y)*-coordinates* (3.34) *for the* α*-simplex.*

(1) For $\beta \leq 2$ *the rate function has a unique global minimum for any* α*.*

(2) For $2 < \beta \leq \frac{18}{7}$ *the rate function has precisely two global minimizers if* α *lies in the segment* $\{0\} \times \left(-\frac{1}{2}, -\frac{1-(\beta-2)e^{3-\beta}}{2+(\beta-2)e^{3-\beta}}\right)$ *or its images under the permutation group. For any other* α *the rate function has a unique global minimum (see Figure 3.4).*

(3) For $\frac{18}{7} < \beta < 4\log 2$ *the rate function has precisely two global minimizers if* α *lies in the segment* $\{0\} \times (-\frac{1}{2}, y_{\text{triple}}(\beta))$ *or if* α *lies on the curve that is a solution to the initial value problem* (3.60) *or its images on the permutation group. Here,* $y_{\text{triple}}(\beta)$ *is the* y*-coordinate of the triple point. If* α *is a triple point, the rate function has three global minimizers. For any other* α *it has a unique global minimum. (see Figures 3.5 and 3.6).*

(4) For $\beta \geq 4\log 2$ the rate function has two global minimizers for any α on the segment $(-\frac{1}{2}, 0) \times \{0\}$ or its images under the permutation group. It has four global minimizers if $\beta = 4\log 2$ and three global minimizers if $\beta > 4\log 2$ for $\alpha = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. For any other α it has a unique global minimum (see Figure 3.7).

Now it is clear from the previous section that for $\beta \leq 2$ we do not see multiple global minimizers because points in the bifurcation set have $\beta > 2$ and in a high temperature regime we have a unique global minimum. The other three regimes are interesting and it is useful to keep the bifurcation set in mind when analysing these. However, before we discuss the other regimes in detail let us state another observation which clarifies the word "surface" of the term *coexistence surface:* Locally (except at the triple points), this set is indeed a two-dimensional submanifold.

Suppose (β, α) is such that the rate function $f_{\beta,\alpha}$ has two distinct non-degenerate stationary points $\nu_{\beta,\alpha}$ and $\mu_{\beta,\alpha}$. Since the rate function depends smoothly on its parameters, the Implicit Function Theorem tells us that we find two smooth maps $(\beta, \alpha) \mapsto \nu_{\beta,\alpha}$ and $(\beta, \alpha) \mapsto \mu_{\beta,\alpha}$ that map a neighbourhood U of (β, α) to Δ^2 such that $\nu_{\beta,\alpha}$ and $\mu_{\beta,\alpha}$ are stationary points of $f_{\beta,\alpha}$ for every (β, α) in the neighbourhood.

Lemma 26. *The set of (β, α) such that $f_{\beta,\alpha}(\nu_{\beta,\alpha}) = f_{\beta,\alpha}(\mu_{\beta,\alpha})$ is a two-dimensional embedded submanifold of $(0, \infty) \times \Delta^2$.*

Proof. Let us define the smooth map F from U to $\mathbb{R}$ via

$$F(\beta, \alpha) = f_{\beta,\alpha}(\nu_{\beta,\alpha}) - f_{\beta,\alpha}(\mu_{\beta,\alpha}). \tag{3.53}$$

We now want to apply the *Constant-Rank Level Set Theorem* [43] for F to conclude the proof. The differential of F in terms of the (β, u, v) coordinates is given via the row vector

$$\Big(-\tfrac{1}{2}(\|\nu_{\beta,\alpha}\|^2 - \|\mu_{\beta,\alpha}\|^2), \quad (\mu_{\beta,\alpha})_1 - (\nu_{\beta,\alpha})_1, \quad (\mu_{\beta,\alpha})_2 - (\nu_{\beta,\alpha})_2\Big) \tag{3.54}$$

which is the zero map if and only if $\nu_{\beta,\alpha} = \mu_{\beta,\alpha}$. Thus, the differential has constant rank one. □

3.3.1 Coexistence in the regime of the rockets

In the regime of the rockets ($2 < \beta \leq \frac{18}{7}$) the only cell that yields two local minimizers is given by the region enclosed by the rockets. The Maxwell set of this region is given by the intersection with the axes of symmetry. This is due to the fact that the two local minimizers lie in different fundamental cells of the simplex and that an asymmetry in the fields α leads to the same asymmetry in the global minimizer. This is explained by the following lemma which is inspired by Lemma 1 of [58].

Lemma 27 (Tilting Lemma)**.** *Let ν be a global minimum of $f_{\beta,\alpha}$. If $\alpha_i > \alpha_j$, then $\nu_i > \nu_j$.*

Proof. Let $\nu'_j = \nu_i$, $\nu'_i = \nu_j$ and $\nu'_k = \nu_k$ for k not in $\{i, j\}$. Since

$$f_{\beta,\alpha}(\nu') - f_{\beta,\alpha}(\nu) = (\nu_i - \nu_j)(\log \alpha_i - \log \alpha_j) \geq 0 \tag{3.55}$$

and $\log \alpha_i - \log \alpha_j > 0$, we conclude $\nu_i \geq \nu_j$. Assume $\nu_i = \nu_j$ and consider the push-forward of the tangent vector $v = \frac{\partial}{\partial \nu_j} - \frac{\partial}{\partial \nu_i}$:

$$df_{\beta,\alpha}(v) = \log \alpha_i - \log \alpha_j > 0. \tag{3.56}$$

Thus ν is not a stationary point which contradicts the fact that ν is a minimizer. Therefore $\nu_i > \nu_j$. □

The cusp point of the rockets is given by the end point of the curve $\chi_\beta \circ \gamma_\beta(\nu_1)$ which in this regime is $\nu_1 = 1 - \frac{2}{\beta}$ (see Theorem 18). Thus the Maxwell set in (x, y)-coordinates is the segment

$$\{0\} \times \left(-\frac{1}{2}, -\frac{1 - (\beta - 2)e^{3-\beta}}{2 + (\beta - 2)e^{3-\beta}} \right). \tag{3.57}$$

3.3.2 Coexistence in the regime of disconnected pentagrams

In the regime $\frac{18}{7} < \beta < \frac{8}{3}$ three pentagrams have already unfolded but are still disconnected. As we have already discussed there are (modulo symmetry) three cells with two local minimizers and one cell with three local minimizers. The Maxwell set in cell I (see Figure 3.5) is the easiest. Here we have again two minimizers in different fundamental cells and

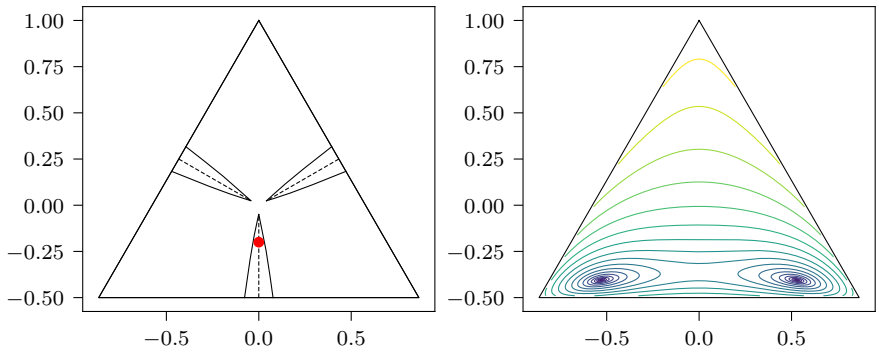

Figure 3.4: The left plot shows the Maxwell set (dashed lines) inside of the rockets. The right plot shows the contours of a typical point on the Maxwell set as indicated by the red dot in the left plot. The minimizers are equally deep and are mapped onto each other under reflection at the vertical axis.

therefore the Maxwell set is the intersection with the respective axis of symmetry. In the cells III and IV we also have two local minimizers but they lie in the same fundamental cell. Therefore the Tilting Lemma (Lemma 27) does not apply and the Maxwell set is a curved line deviating from the axis of symmetry. Cell II is special because we have three different local minimizers two of which lie in different fundamental cells. Since the classical phase diagram describes the degeneracy of the *global* minimum, we know that the Maxwell set continues on the axis of symmetry for as long as the two local minima from the different fundamental cells are lower than the third minimum. There exists however a point on the axis of symmetry at which this behaviour changes: the triple point. This is a point at which all minimizers are global minimizers. Since two of the local minimizers lie in different fundamental cells, the triple point must lie on the "star" (Lemma 27), that is, at least two components are equal. This is also the point where the Maxwell set leaves the axis of symmetry because the minimizers involved do not lie in different fundamental cells anymore. It suffices to compute the Maxwell set in either cell III or IV because of symmetry. The problem of computing the Maxwell set can be transformed into a solution of an initial value problem where the initial value is given by the triple point.

Proposition 28. *For each positive β there exists exactly one point α*

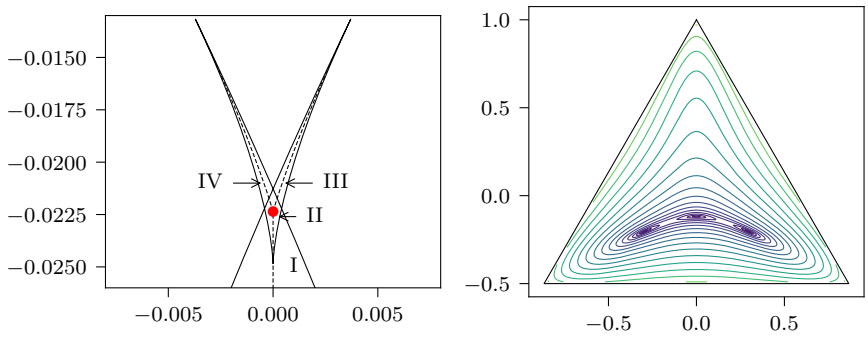

Figure 3.5: The left plot shows a magnification of one of the three pentagrams with its containing Maxwell set (dashed lines). The red dot marks the triple point at which all three minimizers are global minimizers. This can be seen in the right plot which shows a contour plot of the potential at the triple point.

with $\alpha_1 \leq \alpha_2 \leq \alpha_3$ such that $f_{\beta,\alpha}$ has precisely three global minimizers.

Proof. It is clear that the triple point α lies on the axis of symmetry. Therefore, let the curve $u \mapsto \alpha(u)$ be the axis of symmetry intersected with cell II. Then we have two local minimizers $\mu(u)$ and $\nu(u)$ of $f_{\beta,\alpha(u)}$ such that $\mu_1(u) > \nu_1(u)$. The difference

$$g(u) := f_{\beta,\alpha(u)}(\nu(u)) - f_{\beta,\alpha(u)}(\mu(u)) \tag{3.58}$$

is a monotonically increasing function of u:

$$g'(u) = \frac{\partial f_{\beta,\alpha(u)}}{\partial u} = \mu_1(u) - \nu_1(u) > 0. \tag{3.59}$$

□

Proposition 29. *The set of all α such that the two local minimizers inside of the fundamental cell $\nu_1 < \nu_2 < \nu_3$ have the same depth is given in (u, v)-coordinates by the graph of $u \mapsto v(u)$ which is the solution of the initial value problem*

$$\begin{aligned} \frac{dv}{du} &= -\frac{\nu_1(u,v) - \mu_1(u,v)}{\nu_2(u,v) - \mu_2(u,v)} \\ v(u_0) &= v_0 \end{aligned} \tag{3.60}$$

where $\alpha = (u_0, v_0)$ is the triple point, $\nu(u,v)$ and $\mu(u,v)$ are the two local minimizers of $f_{\beta,\alpha}$ in the same cell for $\alpha = (u, v(u))$.

Proof. The curve fulfills

$$f_{\beta,\alpha(u)}(\nu(u,v)) = f_{\beta,\alpha(u)}(\mu(u,v)). \tag{3.61}$$

Note that

$$\frac{d}{du} f_{\beta,\alpha(u)}(\nu(u,v)) = \nu_1(u,v) - \nu_2(u,v)\frac{dv}{du} \tag{3.62}$$

since $\nu(u,v)$ is a stationary point. This proves the proposition. □

However, numerically computing the Maxwell set using this characterization is difficult because the mapping $u \mapsto (\nu(u,v), \mu(u,v))$ is not explicit. Therefore we consider the system of equations

$$\begin{aligned} f_{\beta,\alpha}(\mu) &= f_{\beta,\alpha}(\nu) \\ \chi_\beta(\mu) &= \chi_\beta(\nu) \end{aligned}$$

instead. Because the free energy at a stationary point is given by

$$f_{\beta,\alpha}(\nu) = \sum_{i=1}^{3} \left(\frac{1}{2}\beta\nu_i^2 + \nu_i \log \sum_{j=1}^{3} \nu_j e^{-\beta\nu_j} \right) \tag{3.63}$$

this system does not depend on α and can be solved numerically for fixed values of β and μ_1. For better stability of the numerics we start with the triple point where the minimizers are well separated and then iteratively use the results as initial values for the next numerical step. In this way, the Figures 3.4 to 3.7 were obtained.

3.3.3 From beak-to-beak to Ellis-Wang

The qualitative nature of the Maxwell sets does not change with increasing β after the beak-to-beak scenario until we reach the Ellis-Wang point. The cells with two minima which resulted from the merging of the horns of the pentagrams contain two minimizers which lie in the same fundamental cell as discussed in the previous subsection. Therefore the Maxwell set is again given by the solution to the initial value problem (3.60). This continues even after the crossing temperature where the central cell now

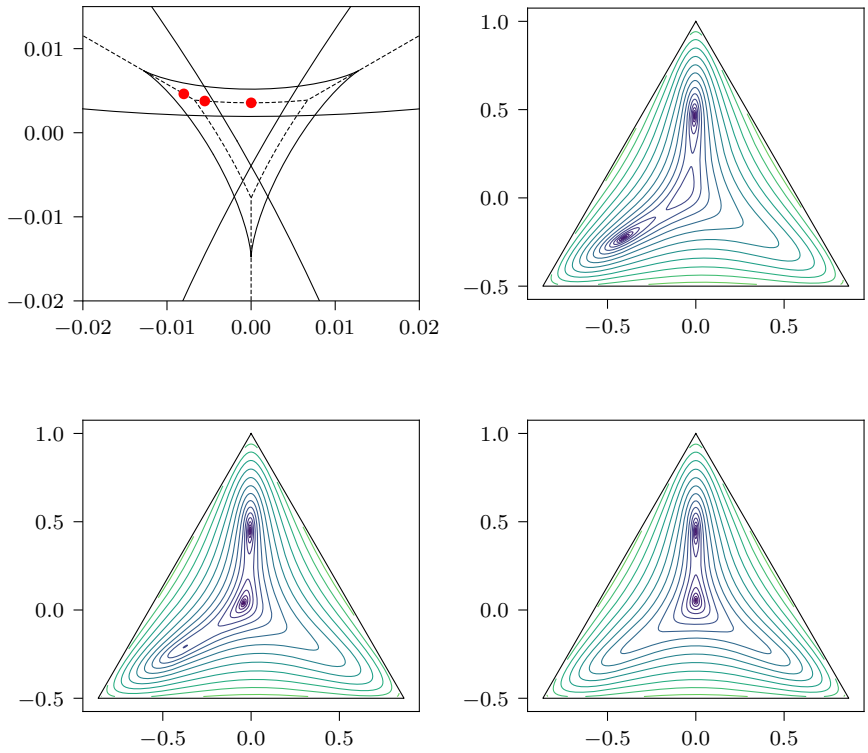

Figure 3.6: The upper left plot shows a magnification of the center of the simplex showing the bifurcation set together with the Maxwell set (dashed lines) for a β between beak-to-beak and Ellis-Wang. Following the red dots on the Maxwell set from left to right we see how the potential changes in the upper right, lower left and lower right plots.

contains four minima. Before the Ellis-Wang temperature the central fourth minimum is a local but not global minimum. The two outer minima each lie in the same fundamental cell so that the initial value problem applies. However, this changes after the Ellis-Wang point.

3.3.4 Beyond Ellis-Wang

At $\beta = 4\log 2$ the outer minima and the central local minimum are equally deep. After the Ellis-Wang point ($\beta > 4\log 2$), the outer minima are lower than the central minimum. They are equally deep in zero

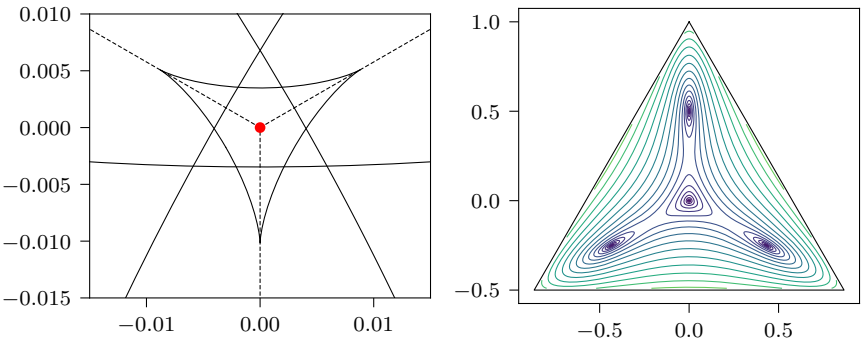

Figure 3.7: The left plot shows the Maxwell set (dashed lines) beyond Ellis-Wang together with the bifurcation set at the Ellis-Wang $\beta = 4 \log 2$. The Maxwell set looks the same at lower temperatures. The potential shows four equally deep minima for α at the red dot. Further decreasing the temperature, the central minimum will raise and eventually become a maximum.

magnetic field and it is still possible using the Tilting Lemma to break the symmetry of the fields partially and achieve two equally deep minimizers. This can be done in the regime where the rate function has a local minimum in the centre ($\beta < 3$) as well as in the regime where the local minimum has become a local maximum ($\beta > 3$).

4 Dynamical Gibbs–non-Gibbs transitions in the Curie–Weiss Potts model in the regime $\beta < 3$

This chapter contains joint work with Prof. Dr. Christof Külske at Ruhr-Universität Bochum [37].

4.1 Introduction

4.1.1 Research context

The past years have seen progress from various directions in the understanding of Gibbs–non-Gibbs transitions for trajectories of measures under time-evolution, and also more general transforms of measures. The Gibbs property of a measure describing the state of a large system in statistical mechanics is related to the continuity of single-site conditional probabilities, considered as a function of the configuration in the conditioning. If a measure becomes non-Gibbsian, there are internal mechanisms which are responsible for the creation of such discontinuous dependence. This leads to the study of *hidden phase transitions*, which was started in the particular context of renormalization group pathologies in van Enter, Fernández, and Sokal [56].

Such studies have been made for a variety of systems in different geometries, for different types of local degrees of freedom, and under different transformations. Let us mention here time-evolved discrete lattice spins [32, 53], continuous lattice spins [39, 57], time-evolved models of point particles in Euclidean space [30], and models on trees [55]. For a discussion of non-Gibbsian behavior of time-evolved lattice measures in regard to the approach to a (possibly non-unique) invariant state under dynamics, see [29], for relevance of non-Gibbsianness to

the infinite-volume Gibbs variational principle (and its possible failure) see [36, 40]. For recent developments for one-dimensional long-range systems, and the relation between continuity of one-sided (vs. two-sided) conditional probabilities see [2–4, 54].

In the present paper we are aiming to contribute to the understanding of Gibbs–non-Gibbs transformations for mean-field models, in the sense of the sequential Gibbs property [10, 15–17, 26, 28, 31, 35]. Usually there is a somewhat incomplete picture for lattice models, due to the difficulty to find sharp critical parameters. Mean-field models on the other hand are often "solvable" in terms of variational principles which arise from the large deviation formalism, while the remaining model-dependent task to characterize the minimizers and understand the corresponding various bifurcations can be quite substantial. We choose to work for our problem in the so-called two-layer approach, in which one needs to understand the parameter dependence of the large-deviation functional of a conditional first-layer system. In this functional the conditioning provides an additional parameter given by an empirical measure on the second layer. This is more direct than working in the Lagrangian formalism on trajectory space, which would provide additional insights on the nature of competing histories that explain the current state of the system at a discontinuity point [15, 34, 50, 52].

Compared to the Curie–Weiss Ising model, the Fuzzy Potts model and the Widom-Rowlinson models, we find in the present analysis of the time-evolved Curie–Weiss Potts model significantly more complex transition phenomena, see Theorem 31 and Figure 4.2. This has to be expected as already the behavior of the fully non-symmetric static model is subtle [38]. It forces us to make use of the computer for exact symbolic computations, in the derivation of the transition curves (BU, ACE and TPE in Figure 4.2, discussed in Sects. 4.4.4, 4.4.5 and 4.4.6), along with some numerics for our bifurcation analysis. We believe that these tools (see page 107) may also be useful elsewhere.

Now, our approach rests on singularity theory [1, 6, 20, 21, 49] for the appropriate conditional rate functional of the dynamical model. This provides us with a four-parameter family of potentials, for a two-dimensional state-variable taking values in a simplex. It turns out that the understanding of the parameter dependence of the dynamical model is necessarily based on the good understanding of the bifurcation geometry

of the free energy landscape of the static case for general vector-valued fields [38]. In that paper, which generalizes the results of Ellis and Wang [14] and Wang [58], we lay out the basic methodology. Therein we also explain the phenomenology of transitions (umbilics, butterflies, beak-to-beak) from which we need to build here for the dynamical problem.

As a result of the present paper we show that the unfoldings of the static model indeed reappear in the dynamical setup, and acquire new relevance as hidden phase transitions. It is important to note that, in order for this to be true, we have to restrict to mid-range inverse temperatures $\beta < 3$. More work has still to be done to treat the full range of inverse temperatures for the dynamical model, where more general transitions seem to appear for very low temperatures. For the scope of the present paper, it is this close connection between the static model [38] in fully non-symmetric external fields, and the symmetrically time-evolved symmetric model in intermediate β range, which is really crucial to unravel the types of trajectories of bad empirical measures of Theorem 31. It would be challenging to exploit whether an analogous non-trivial connection, that we observe for our particular model, holds for more general classes of models. This clearly asks for more research.

4.1.2 Overview and organization of the paper

In the present paper we study the simplest model which is, together with its time-evolution, invariant under the permutation group with three elements: We consider the 3-state Curie–Weiss Potts model in zero external field, under an independent symmetric stochastic spin-flip dynamics. Based on previous examples [35], one may expect loss without recovery of the Gibbs property for all initial temperatures lower than a critical one (which then may or may not coincide with the critical temperature of the initial model), and Gibbsian behavior for all times above the same critical temperature. We show that this is not the case for our model, and the behavior is much more complicated: The trajectories of the model show a much greater variety, depending on the initial temperature. We find a regime of Gibbs forever (I), a regime of loss with recovery (II) and a regime of loss without recovery (III). Figure 4.1 shows the non-Gibbs region in the two-dimensional space of initial temperature and time. The boundary of this non-Gibbs region

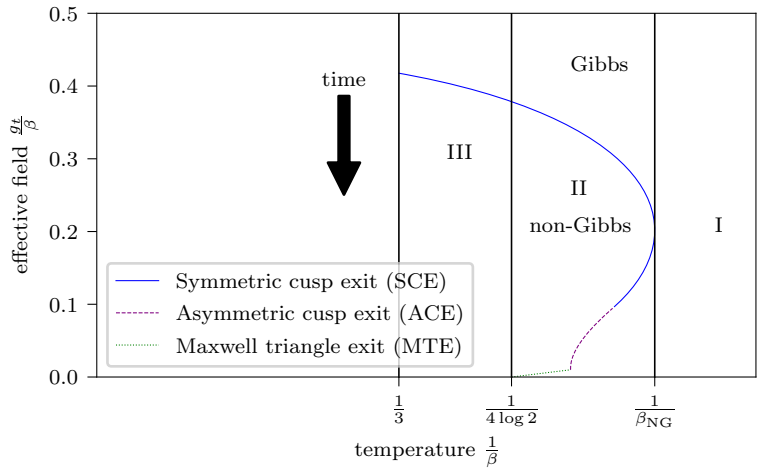

Figure 4.1: This figure shows the non-Gibbs region for the mid-range temperature regime we consider. The boundary of this region consists of three different curves which correspond to exit scenarios of bad empirical measures.

consists of three different curves which correspond to *exit scenarios* of different types of *bad empirical measures*. Bad empirical measures are points of discontinuity of the limiting conditional probabilities as defined in Definition 30. Under the time evolution $t \uparrow \infty$ (or equivalently $g_t \downarrow 0$ given by (4.4)) the system moves along vertical lines of fixed β towards the temperature axis. Intersections with a finite number of lines occur along this way, which are responsible for the transitions described in our main theorem, Theorem 31. These additional relevant lines are shown in Figure 4.2. Theorem 31 rests on the understanding of the structure of stationary points of the time-dependent conditional rate function given in Formula (4.9) via singularity theory.

It turns out that the bifurcations we encounter for general values of the four-dimensional parameter $(\alpha, \beta, t) \in \Delta^2 \times (0, \infty) \times (0, \infty)$ (see (4.6)) are of the same types as for the static model depending on a three-dimensional parameter. However, this holds only *if* we restrict to mid-range inverse temperatures $\beta < 3$ *and* to endconditionings α taking values in the unit simplex (and not in the full hyperplane spanned by the simplex). Nevertheless, in order to understand the relevant singularities, the analysis is best done by first relaxing the probability

measure constraint on the parameter α and allow it to take values in the hyperplane. The analysis proceeds with a description of the *bifurcation set*, where the structure of stationary points of the conditional rate function changes, and the *Maxwell set*, where multiple global minimizers appear. To pick from these transitions the ones which are relevant to the problem of sequential Gibbsianness and visible on the level of bad empirical measures, we have to take the probability measure constraint for α into account. This step is neither necessary in the static Potts nor in the dynamical symmetric Ising model. The lines *Symmetric cusp exit (SCE), Asymmetric cusp exit (ACE), Triple point exit (TPE) and Maxwell triangle exit (MTE)* depicted in the full phase diagram in Figure 4.2 are examples of such exit scenarios. For those lines there is an exit of a certain particular critical value of α from the unit simplex (observation window). The detailed dynamical phase diagram in Figure 4.2 shows more information about the transitions during time evolution. Preliminary investigations show that the structural similarity with the static case may no longer be valid in the regime $\beta > 3$. Therefore we leave the region of very low temperatures for future research.

We describe the model we are considering together with its time-evolution in Sect. 4.1.3 where we also define what we mean by Gibbsianness (or the sequential Gibbs property). In Sect. 4.2 we present our main theorem and describe the transitions of the sets of bad empirical measures as a function of the parameters β and t. We will establish the connection between the analysis of the potential function $G_{\alpha,\beta,t}$ and the Gibbs property of the time-evolved model in Sect. 4.3. The analysis of the potential function using the methods of singularity theory is then carried out in the Sects. 4.4 and 4.5.

4.1.3 The model and sequential Gibbsianness

We consider the mean-field Potts model with three states in vanishing external field under an independent symmetric spin-flip dynamics. The space of configurations in finite-volume $n \geq 2$ is defined as $\Omega_n = \{1, 2, 3\}^n$ and the Hamiltonian of the initial model is

$$H_n(\sigma) = -\frac{1}{2n} \sum_{i,j=1}^{n} \delta_{\sigma_i,\sigma_j}. \tag{4.1}$$

So at time $t = 0$ the distribution of the model is given by

$$\mu_{n,\beta}(\sigma) = \frac{e^{-\beta H_n(\sigma)}}{\sum_{\tilde{\sigma}\in\Omega_n} e^{-\beta H_n(\tilde{\sigma})}}. \tag{4.2}$$

We consider a rate-one symmetric spin-flip time-evolution in terms of independent Markov chains on the sites with transition probabilities

$$p_t(a,b) = \frac{e^{g_t 1_{b=a}}}{e^{g_t}+2} \tag{4.3}$$

from state a to b where

$$g_t = \log\frac{1+2e^{-3t}}{1-e^{-3t}}. \tag{4.4}$$

We are interested in the Gibbsian behavior of the time-evolved measure

$$\mu_{n,\beta,t}(\eta) = \sum_{\sigma\in\Omega_n} \mu_{n,\beta}(\sigma)\prod_{i=1}^{n} p_t(\sigma_i,\eta_i). \tag{4.5}$$

The *unit simplex*

$$\Delta^2 = \{\nu\in\mathbb{R}^3 \mid \nu_i \geq 0, \sum_{i=1}^{3}\nu_i = 1\} \tag{4.6}$$

contains the empirical distributions of spins. By Gibbsian behavior we mean the existence of limiting conditional probabilities in the following sense.

Definition 30. *The point α in Δ^2 is called a* good point *if and only if the limit*

$$\gamma_{\beta,t}(\cdot|\alpha) := \lim_{n\to\infty}\mu_{n,\beta,t}(\cdot|\eta_{n,2},\ldots,\eta_{n,n}) \tag{4.7}$$

exists for every family $\eta_{n,k}\in\{1,2,3\}$ with $n\geq 2$ and $2\leq k\leq n$ such that

$$\lim_{n\to\infty}\frac{1}{n-1}\sum_{k=2}^{n}\eta_{n,k} = \alpha. \tag{4.8}$$

We call α bad, if it is not good. The model $\mu_{\beta,t}$ is called sequentially Gibbs *if all α in the unit simplex Δ^2 are good points.*

4.2 Dynamical Gibbs–non-Gibbs transitions: main result

Our main result on the dynamical Gibbs–non-Gibbs transitions in the high-to-intermediate temperature regime for the initial inverse temperature $\beta < 3$ is as follows. This temperature regime ranges from high temperature, covering the phase transition temperature (Ellis-Wang inverse temperature $\beta = 4 \log 2$), up to the elliptic umbilic point $\beta = 3$ (where the central stationary point of the time-zero rate function in zero external field changes from minimum to maximum).

Essential parts of the structure of the trajectories of dynamical transitions as a function of time t in the regime $\beta < 3$ remain unchanged over the three inverse-temperature intervals I, II and III, which were already visualized in Figure 4.1. The type of transitions can be understood as deformations of the sequences of transitions found in the static Potts model in general vector-valued fields analyzed in [38], where in that case only the one-dimensional parameter β was varied. Observe that however, the dynamical transitions we describe here, do not necessarily occur in a monotonic order with respect to what is seen in the static model under temperature variation. This is for instance (but not only) apparent in the phenomenon of recovery of Gibbsianness. At very low temperatures ($\beta > 3$) different bifurcations seem to occur which will be left for future research. While reading the following theorem it is useful to have Figure 4.2 in mind as the inverse temperatures and transition times are related to the lines depicted in the *dynamical phase diagram*.

Theorem 31. *Consider the time-evolved Curie–Weiss Potts model given by (4.1–4.2) in zero external field, for initial inverse temperature $\beta > 0$ and at time $t > 0$ under the symmetric spin-flip dynamics (4.3–4.5). Then the following holds.*

(I) For $\beta < \beta_{\mathrm{NG}} \approx 2.52885$ the time-evolved model is sequentially Gibbs for all $t > 0$.

(II) For $\beta_{\mathrm{NG}} < \beta < 4 \log 2$ the time-evolved model loses and then recovers the Gibbs property at sharp transition times. More precisely, there exist $\beta_{\mathrm{BE}} < \beta_*$ in this interval such that the following types of trajectories of sets of bad empirical measures occur:

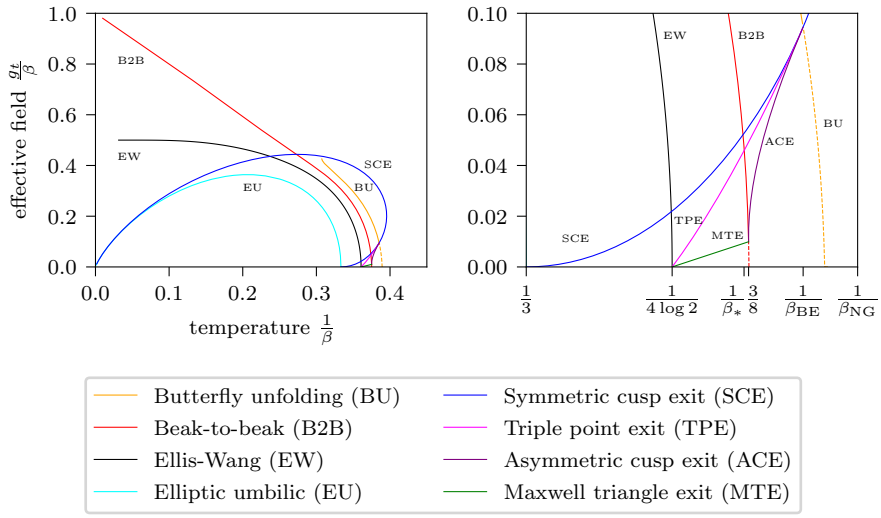

Figure 4.2: This figure shows the dynamical phase diagram which displays all lines in the two-dimensional space of $\frac{1}{\beta}$ and $\frac{g_t}{\beta}$ at which the structure of the bifurcation set slice or the Maxwell set slices changes. We have also marked the six important temperatures in the magnified plot on the right.

(i) *For $\beta < \beta_{\mathrm{BE}}$ the bad empirical measures are given by three symmetric straight lines which are first growing with time from the midpoints of the simplex edges towards the center, then shrinking with time again.*

(ii) *For $\beta_{\mathrm{BE}} < \beta < \frac{8}{3}$ the bad empirical measures are given by three symmetric straight lines in a first time interval $t_{\mathrm{NG}}(\beta) < t < t_{\mathrm{BU}}(\beta)$. For a second time interval $t_{\mathrm{BU}}(\beta) < t < t_{\mathrm{TPE}}(\beta)$, the set of bad empirical measures consists of three symmetric Y-shaped sets not touching. For $t_{\mathrm{TPE}}(\beta) < t < t_{\mathrm{ACE}}(\beta)$ the set of bad empirical measures consists of six disconnected arcs. For $t > t_{\mathrm{ACE}}(\beta)$ the system is Gibbsian again.*

(iii) *For $\frac{8}{3} < \beta < \beta_*$ and $t_{\mathrm{NG}}(\beta) < t < t_{\mathrm{BU}}(\beta)$ the bad empirical measures consist of three symmetric straight lines. For $t_{\mathrm{BU}}(\beta) < t < t_{\mathrm{TPE}}(\beta)$, the set of bad empirical measures consists of three Y-shaped sets not touching. For $t_{\mathrm{TPE}}(\beta) < t < t_{\mathrm{B2B}}(\beta)$ the set of bad empirical measures consists of six*

disconnected arcs. For $t_{\mathrm{B2B}}(\beta) < t < t_{\mathrm{MTE}}(\beta)$ *the set of bad empirical measures consists of three disconnected arcs. For* $t > t_{\mathrm{MTE}}(\beta)$ *the system is Gibbsian again. The inverse temperature* β_* *is given by the intersection point of the two lines B2B and TPE in Figure 4.2.*

(iv) For $\beta_* < \beta < 4\log 2$ *and* $t_{\mathrm{NG}}(\beta) < t < t_{\mathrm{BU}}(\beta)$ *the bad empirical measures consist of three symmetric straight lines. For* $t_{\mathrm{BU}}(\beta) < t < t_{\mathrm{B2B}}(\beta)$*, the set of bad empirical measures consists of three Y-shaped sets not touching. For* $t_{\mathrm{B2B}}(\beta) < t < t_{\mathrm{TPE}}(\beta)$ *the set of bad empirical measures consists of a triangle with curved edges and three symmetric straight lines attached. For* $t_{\mathrm{TPE}}(\beta) < t < t_{\mathrm{MTE}}(\beta)$ *the set of bad empirical measures consists of three disconnected arcs. For* $t > t_{\mathrm{MTE}}(\beta)$ *the system is Gibbsian again.*

(III) For $4\log 2 < \beta < 3$ *the time-evolved model loses the Gibbs property without recovery at a sharp transition time and the set of bad empirical measures has the following structure: For* $t \le t_{NG}(\beta)$ *the time-evolved model is Gibbsian. For* $t_{NG}(\beta) < t < t_{\mathrm{BU}}(\beta)$ *the bad empirical measures are given by three symmetric straight lines which are growing with time and become Y-shaped sets for* $t_{\mathrm{BU}}(\beta) < t < t_{\mathrm{B2B}}(\beta)$*. For* $t_{\mathrm{B2B}}(\beta) < t < t_{\mathrm{EW}}(\beta)$ *the sets then touch and form one connected component consisting of a central triangle with three straight lines attached to the vertices. The central triangle then shrinks to a point at* $t = t_{\mathrm{EW}}(\beta)$ *and the bad empirical measures are given by three symmetric straight lines which meet in the simplex center for all* $t > t_{\mathrm{EW}}(\beta)$*.*

The meaning and computation of these lines are discussed in Sects. 4.4 and 4.5. While only the three lines SCE, ACE and MTE appear as part of the boundary line of the non-Gibbs region, the other lines are relevant for structural changes of the set of bad empirical measures. There are lines which are explicit in the sense that they are given in terms of zeros of one-dimensional non-linear functions, for example, the entry time $t_{\mathrm{NG}}(\beta)$ (formula (4.60)) or the butterfly unfolding time $t_{\mathrm{BU}}(\beta)$ (Formula (4.72)). The least explicit lines are the MTE and TPE lines which involve a Maxwell set computation, the most explicit line is SCE which is given

in parametric form $s \mapsto (\beta(s), g_t(s))$ as described in Proposition 37. Figure 4.3 gives a graphical overview of the possible types of sequences of bad empirical measures with increasing time for the different temperature regimes. There is an even more detailed graphic that illustrates all the transitions involved in the bifurcation set as well as in the Maxwell set. You can find this graphic in the electronic supplemental material (ESM) under the filename `detailed_overview.pdf`.

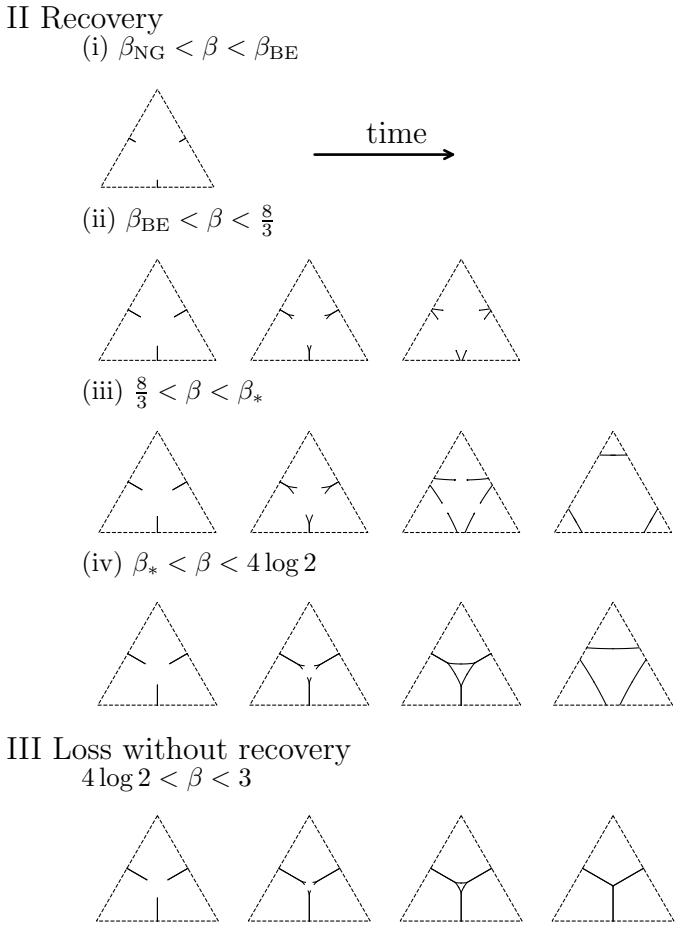

Figure 4.3: These are the typical sequences of bad empirical measures α for the inverse temperature regimes described in Theorem 31. With increasing time, you can observe the structural change of the set of bad empirical measures as it passes the various transition times. For example in (II.ii) straight lines enter the simplex, become non-touching Y-shaped sets at the butterfly transition time $t_{\mathrm{BU}}(\beta)$ and move out of the simplex. The midpoints of the Y-shaped sets exit at $t_{\mathrm{TPE}}(\beta)$ and the set leaves the simplex completely at $t_{\mathrm{ACE}}(\beta)$. In (II.iii) the midpoints of the Y-shaped sets leave the unit simplex at $t_{\mathrm{TPE}}(\beta)$ and the two respective arcs connect at the beak-to-beak transition time $t_{\mathrm{B2B}}(\beta)$. The remaining three arcs move towards the corners and leave the unit simplex at $t_{\mathrm{MTE}}(\beta)$. The exit of the midpoints of the Y-shaped sets and the connection of the six arcs occurs in reversed order in the next row (II.iv). In (III) the central triangle shrinks to a point and forms the star-like set that remains in the simplex forever.

4.3 Infinite-volume limit of conditional probabilities

The existence of the infinite-volume limit of the conditional probabilities, that is, the question of sequential Gibbsianness, can be transformed into an optimization problem of a certain potential function. As the parameters (β, t) are fixed throughout this section let us write μ_n for the measure $\mu_{n,\beta,t}$.

Theorem 32. *Suppose the function $G_{\alpha,\beta,t}\colon \mathbb{R}^3 \to \mathbb{R}$ (known as Hubbard-Stratonovič (HS) transform) given by*

$$G_{\alpha,\beta,t}(m) = \frac{1}{2}\beta\langle m, m\rangle - \sum_{b=1}^{3} \alpha_b \log \sum_{a=1}^{3} e^{\beta m_a + g_t 1_{a=b}} \tag{4.9}$$

has a unique global minimizer, then α is a good point, that is, the infinite-volume limit of the conditional probabilities $\mu_n(\cdot|\alpha_n)$ with $\alpha_n \to \alpha$ exists independently of the choice of (α_n).

The idea of the proof goes as follows: We can rewrite the conditional probabilities $\mu_n(\cdot|\alpha_n)$ in terms of an expected value with respect to a disordered mean-field Potts model $\bar{\mu}_n$ (see Lemma 33). Thus, we have to study the weak convergence of L_n, where L_n is the empirical distribution of the spins $\sigma_2, \dots, \sigma_n$. Note that this is equivalent to the weak convergence of $\frac{W}{\sqrt{\beta(n-1)}} + L_n$ with some independent standard normal variable W. Because of the representation of the distribution of $\frac{W}{\sqrt{\beta(n-1)}} + L_n$ in terms of the function $G_{\alpha_n,\beta,t}$ (Lemma 34), we can prove the theorem by an asymptotic analysis of integrals of the form

$$\int_{\mathbb{R}^3} f(m) e^{-(n-1)G_{\alpha_n,\beta,t}(m)} \, \mathrm{d}m \tag{4.10}$$

as was done by Ellis and Wang [14]. So it suffices to prove the Lemmata 33 and 34. A point is good if the respective random field model shows no phase transition, that is, the law of large numbers holds. To be precise, we have the following representation:

Lemma 33. *The finite-volume conditional probabilities are given by*

$$\mu_n(\eta_1|\eta_2, \dots, \eta_n) = \left\{ \sum_{\tilde{\eta}_1} \frac{\int f_n^{\tilde{\eta}_1} \, \mathrm{d}\bar{\mu}_n[\eta_2, \dots, \eta_n]}{\int f_n^{\eta_1} \, \mathrm{d}\bar{\mu}_n[\eta_2, \dots, \eta_n]} \right\}^{-1} \tag{4.11}$$

where the observable $f_n^{\eta_1}$ *is defined via*

$$f_n^{\eta_1}(\sigma_2,\dots,\sigma_n) = \sum_a \exp\left(\frac{\beta}{n}\sum_{i=2}^{n} 1_{\sigma_i=a}\right) p_t(a,\eta_1) \tag{4.12}$$

and the constrained first-layer model $\bar{\mu}_n$ *is given by*

$$\bar{\mu}_n[\eta_2,\dots,\eta_n](\sigma_2,\dots,\sigma_n) = \frac{e^{\frac{\beta}{2n}\sum_{i,j=2}^{n} 1_{\sigma_i=\sigma_j}} \prod_{i=2}^{n} p_t(\sigma_i,\eta_i)}{\sum_{\tilde{\sigma}_2,\dots,\tilde{\sigma}_n} e^{\frac{\beta}{2n}\sum_{i,j=2}^{n} 1_{\tilde{\sigma}_i=\tilde{\sigma}_j}} \prod_{i=2}^{n} p_t(\tilde{\sigma}_i,\eta_i)}. \tag{4.13}$$

This is a corrected version of [37, Lemma 4] which has been submitted for publication in the Journal of Statistical Physics.

Proof. The proof is given on page 14 and essentially follows from explicit computations with conditional probabilities. □

This representation of the conditional probabilities transforms the problem of understanding bad points to the analysis of disordered mean-field models and their phase transitions. This analysis is done using the so-called Hubbard-Stratonovič transformation which is successfully used for many models [13, 14, 35].

Lemma 34. *Write*

$$L_n = \frac{1}{n-1}\sum_{i=2}^{n} \delta_{\sigma_i} \tag{4.14}$$

for the empirical measure of $n-1$ *spins with law* $\bar{\mu}_n[\eta_2,\dots,\eta_n]\circ L_n^{-1}$. *Furthermore, let* W *be a standard normal random vector independent of* L_n. *The distribution of* $W/\sqrt{\beta(n-1)} + L_n$ *has a density proportional to* $e^{-(n-1)G_{\alpha_n,\beta,t}}$ *with respect to Lebesgue measure.*

Proof. Denote by $\sigma_2,\dots,\sigma_n$ independent $\{1,2,3\}$-valued random variables each distributed according to $p_t(\mathrm{d}\sigma_i\,,\eta_i)$ with a fixed boundary configuration $\eta_2,\dots,\eta_n$ with empirical measure α_n. We denote the expectation with respect to this distribution by $\mathbb{E}$. Then in order to calculate the distribution of

$$\frac{W}{\sqrt{\beta(n-1)}} + L_n = \frac{W}{\sqrt{\beta(n-1)}} + \frac{1}{n-1}\sum_{i=2}^{n} \delta_{\sigma_i} \tag{4.15}$$

we calculate for every bounded continuous function f the expectation

$$\frac{(2\pi)^{-\frac{3}{2}}}{Z_n}\mathbb{E}\left[\int f\left(w/\sqrt{\beta(n-1)}+L_n\right)e^{-\frac{\|w\|^2}{2}+\frac{\beta}{2}(n-1)\|L_n\|^2}\,\mathrm{d}w\right] \tag{4.16}$$

Now we apply the transformation $m = w/\sqrt{\beta(n-1)} + L_n$ and obtain

$$\frac{(2\pi)^{-\frac{3}{2}}}{Z_n}\mathbb{E}\left[\int f(m)\exp\left(-(n-1)\frac{\beta}{2}\|m\|^2+(n-1)\beta\langle m,L_n\rangle\right)\mathrm{d}m\right] \tag{4.17}$$

In order to complete the proof, we have to calculate the expectation

$$\begin{aligned}\mathbb{E}[\exp((n-1)\beta\langle m,L_n\rangle)] &= \prod_{i=2}^{n}\mathbb{E}[\exp(\beta m_{\sigma_i})]\\ &= \prod_{i=2}^{n}\sum_{a=1}^{3}\frac{e^{\beta m_a+g_t1_{\eta_i=a}}}{e^{g_t}+2}\\ &= \frac{1}{(e^{g_t}+2)^{n-1}}\prod_{i=2}^{n}\sum_{a=1}^{3}e^{\beta m_a+g_t1_{\eta_i=a}}\end{aligned} \tag{4.18}$$

Now we take the logarithm to raise the expression back into the exponent again. So the expected value (4.16) of the bounded continuous function f is equal to the following up to a normalizing constant:

$$\int f(m)\exp\left(-(n-1)\frac{\beta}{2}\|m\|^2+\sum_{i=2}^{n}\log\sum_{a}e^{\beta m_a+g1_{\eta_i=a}}\right)\mathrm{d}m \tag{4.19}$$

We can now identify $G_{\alpha_n,\beta,t}$ in the exponent using that

$$\begin{aligned}\sum_{i=2}^{n}\log\sum_{a=1}^{3}e^{\beta m_a+g_t1_{\eta_i=a}} &= (n-1)\sum_{b=1}^{3}\frac{1}{n-1}\sum_{i=2}^{n}1_{\eta_i=b}\log\sum_{a=1}^{3}e^{\beta m_a+g_t1_{b=a}}\\ &= (n-1)\sum_{b=1}^{3}\alpha_n(b)\log\sum_{a=1}^{3}e^{\beta m_a+g_t1_{b=a}}.\end{aligned} \tag{4.20}$$

□

4.4 Recovery of the Gibbs property

The regime $\beta < \frac{8}{3}$ is split into three parts given by the intervals $(0, \beta_{\mathrm{NG}}]$, $(\beta_{\mathrm{NG}}, \beta_{\mathrm{BE}}]$ and $(\beta_{\mathrm{BE}}, \frac{8}{3})$. In the first part we find that the model is sequentially Gibbs for all times $t > 0$ whereas in the other two parts the system recovers from a state of non-Gibbsian behavior. The driving mechanism in this "recovery regime" is due to the butterfly singularity which is already found in the static model [see 38, Sect. 2.4.1]. However, in contrast to the static model the bifurcation set might leave the unit simplex so that in order to answer the Gibbs–non-Gibbs question the location of this set (and the contained Maxwell set) with respect to the unit simplex is also important.

4.4.1 Elements from singularity theory

In order to investigate the Gibbs–non-Gibbs transitions we have to study the global minimizers of the potential $G_{\alpha,\beta,t}$ (Theorem 32). We will use concepts from singularity theory to derive and explain our results.

Singularity theory allows us to understand how the stationary points of the potential change with varying parameters. This can be achieved by looking at the geometry of the so-called *catastrophe manifold*, which contains the information about the stationary points of the potential for every possible choice of parameter values. More precisely, it consists of the tuples (m, α, β, t) in $\mathbb{R}^3 \times \Delta^2 \times (0, \infty) \times (0, \infty)$ such that m is a stationary point of $G_{\alpha,\beta,t}$ given by (4.9). The *bifurcation set* consists of those parameter values (α, β, t) in $\Delta^2 \times (0, \infty) \times (0, \infty)$ such that there exists a degenerate stationary point m in $\mathbb{R}^3$, that is, a point at which the Hessian has a zero eigenvalue. The parameter values of the bifurcation set give rise to a partition of the parameter space whose cells contain parameters at which the number and nature of stationary points do not change. Although we are only interested in α that are bad *empirical* measures, hence *probability* measures, it is convenient to loosen this constraint and consider α in the hyperplane $H = \{m \in \mathbb{R}^3 | m_1 + m_2 + m_3 = 1\}$ into which the unit simplex is embedded. The following proposition is the basis for the analysis of the bifurcation set.

Proposition 35. *Let Γ denote the map from $\mathbb{R}^3 \times (0,\infty)$ to the space of 3×3 matrices with real entries $\mathrm{Mat}(3,\mathbb{R})$ given by its components*

$$\Gamma_{b,a}(M,t) = \frac{e^{M_a + g_t 1_{b=a}}}{\sum_{c=1}^{3} e^{M_c + g_t 1_{b=c}}}. \tag{4.21}$$

Then we have the following:

(a) Let ρ be any permutation of $\{1,2,3\}$. Then

$$\rho^{-1}\Gamma(M,t)\rho = \Gamma(\rho M, t) \tag{4.22}$$

where we interpret the permutation ρ as a 3×3-matrix and M as a column vector. For example, if $M_2 = M_3$, we find $\Gamma_{3,3}(M,t) = \Gamma_{2,2}(M,t)$ and also $\Gamma_{1,2}(M,t) = \Gamma_{1,3}(M,t)$.

(b) Γ maps $\mathbb{R}^3 \times (0,\infty)$ into the general linear group $\mathrm{GL}(3,\mathbb{R})$ and the inverse matrix of $\Gamma(M,t)$ is given by the formulas

$$\Gamma^{-1}_{a,a}(M,t) = \frac{(e^{g_t}+1)e^{-M_a}}{e^{2g_t}+e^{g_t}-2} \sum_{c=1}^{3} e^{M_c + g_t 1_{c=a}} \tag{4.23}$$

$$\Gamma^{-1}_{b,a}(M,t) = -\frac{e^{-M_b}}{e^{2g_t}+e^{g_t}-2} \sum_{c=1}^{3} e^{M_c + g_t 1_{c=a}} \tag{4.24}$$

for two distinct elements a, b of $\{1,2,3\}$.

(c) The catastrophe manifold of the HS transform $G_{\alpha,\beta,t}$ is the graph of the map $(m,\beta,t) \mapsto \alpha = \chi(m,\beta,t)$ given by

$$\chi(m,\beta,t) = \left(\sum_a m_a \Gamma^{-1}_{a,b}(\beta m, t) \right)_{b=1}^{3} \tag{4.25}$$

from $H \times (0,\infty) \times (0,\infty)$ to H. For $\chi(m,\beta,t)$ to lie in the unit simplex Δ^2 it is necessary (but generally not sufficient) that m lies in Δ^2.

(d) Consider the coordinates $(x, y, z) = \varphi_\beta(m)$ *where*

$$\varphi_\beta(m) = \frac{\beta}{6} \begin{pmatrix} \sqrt{3}(m_3 - m_2) \\ 2m_1 - m_2 - m_3 \\ 2m_1 + 2m_2 + 2m_3 - 2 \end{pmatrix} \tag{4.26}$$

for $m \in \mathbb{R}^3$*. In these coordinates, the* β*-scaled simplex* $\beta\Delta^2$ *is an equilateral triangle in the* (x, y)*-plane centered at the origin. The Hessian matrix* $G''_{\alpha,\beta,t}(m)$ *in these coordinates is in block diagonal form:*

$$\begin{pmatrix} \frac{\partial^2 G_{\alpha,\beta,t}}{\partial x^2} & \frac{\partial^2 G_{\alpha,\beta,t}}{\partial x \partial y} & 0 \\ \frac{\partial^2 G_{\alpha,\beta,t}}{\partial x \partial y} & \frac{\partial^2 G_{\alpha,\beta,t}}{\partial y^2} & 0 \\ 0 & 0 & \frac{3}{\beta} \end{pmatrix} \tag{4.27}$$

The set of degenerate stationary points is given by the solutions (m, β, t) *of the following equation:*

$$\frac{\partial^2 G_{\chi(m,\beta,t),\beta,t}}{\partial x^2} \frac{\partial^2 G_{\chi(m,\beta,t),\beta,t}}{\partial y^2} - \left(\frac{\partial^2 G_{\chi(m,\beta,t),\beta,t}}{\partial x \partial y} \right)^2 = 0 \tag{4.28}$$

Before we present the proof, let us stress the importance of this proposition. The matrix Γ naturally appears in the derivatives of $G_{\alpha,\beta,t}$ and has the two important properties: Firstly, the rows of Γ are probability vectors and secondly the map $M \mapsto \Gamma(M, t)$ is compatible with the symmetry of the model. The fact that the catastrophe manifold is given as a graph allows us to write the bifurcation set as the set of $(\chi(m, \beta, t), \beta, t)$ such that

$$\det G''_{\chi(m,\beta,t),\beta,t}(m) = 0 \tag{4.29}$$

with $(m, \beta, t) \in H \times (0, \infty) \times (0, \infty)$. We can therefore take the same point of view as in the static case [cf. 38, Lemma 3]: We study the zeros of the Hessian determinant as a function of m with β and t fixed. This is a two-dimensional problem since we only have to consider points in the unit simplex Δ^2. Additionally, Δ^2 is bounded so that we can simply compute the zeros of the Hessian determinant numerically on a discretization of Δ^2 as accurately as we want to. In this way we can get insight into the global shape of the bifurcation set. It is convenient to look at this set as composed of the *bifurcation set slices* $B(\beta, t)$, that

is, the subsets for which the parameter (β, t) is fixed. Figure 4.4 shows an example of the zeros of the Hessian determinant together with the respective image under the map $\chi(\cdot, \beta, t)$ for a fixed pair (β, t). We now continue with the proof of the above proposition.

Proof of Proposition 35. Let us prove the claims in increasing order. Fix arbitrary $M \in \mathbb{R}^3$ and positive t. The following equation proves (4.22).

$$\Gamma_{b,a}(\rho M, t) = \frac{e^{M_{\rho(a)} + g_t 1_{b=a}}}{\sum_{c=1}^{3} e^{M_c + g_t 1_{b=c}}} = \frac{e^{M_{\rho(a)} + g_t 1_{\rho(b)=\rho(a)}}}{\sum_{c=1}^{3} e^{M_c + g_t 1_{\rho(b)=c}}} = \Gamma_{\rho(b),\rho(a)}(M, t) \tag{4.30}$$

We proceed with the second point. Note that the matrix $\Gamma(M, t)$ can be written as the product DE of the diagonal matrix $D = (D_{a,b})$ with entries

$$\frac{1_{a=b}}{\sum_{c=1}^{3} e^{M_c + g_t 1_{c=b}}} \tag{4.31}$$

for $a, b \in \{1, 2, 3\}$ and the matrix

$$E = \begin{pmatrix} e^{M_1+g_t} & e^{M_2} & e^{M_3} \\ e^{M_1} & e^{M_2+g_t} & e^{M_3} \\ e^{M_1} & e^{M_2} & e^{M_3+g_t} \end{pmatrix}. \tag{4.32}$$

Since $\det \Gamma(M, t) = \det(D) \cdot \det(E)$ and the determinant of D is clearly positive, we have to check that $\det(E)$ is positive to see that $\Gamma(M, t)$ is in the general linear group. We find that the determinant of E is given by

$$\det(E) = e^{M_1+M_2+M_3}(e^{3g_t} - 3e^{g_t} + 2) \tag{4.33}$$

which is clearly positive for all positive g_t.

To prove the formula for the inverse, let a, b and d be pairwise different elements of $\{1, 2, 3\}$. Substituting the right-hand sides of (4.23–4.24), we

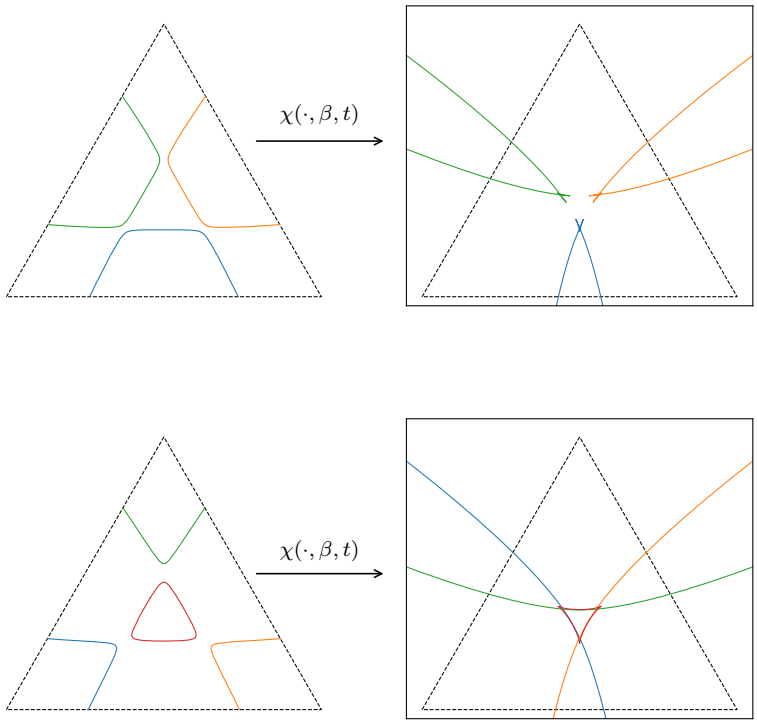

Figure 4.4: The left column shows the solutions to the degeneracy condition (4.28) for $\beta = 2.755$, $g_t = 0.5$ (above) and $g_t = 0.45$ (below) computed using a uniform triangular grid. The right column shows the image of the solutions under the catastrophe map $\chi(\cdot, \beta, t)$ restricted to a square. The branches of the degenerate points on the left and their corresponding images under $\chi(\cdot, \beta, t)$ on the right are marked with the same color. Note that despite the fact that the degenerate stationary points in the left plot lie inside of Δ^2 in the right plot we see that parts of the bifurcation set slice lie outside of the simplex. This is a major difference to the static case.

have the following

$$\begin{aligned}
\Gamma_{b,a}\Gamma^{-1}_{a,a} &= \frac{e^{g_t}+1}{e^{2g_t}+e^{g_t}-2}\frac{\sum_c e^{M_c+g_t 1_{c=a}}}{\sum_c e^{M_c+g_t 1_{c=b}}}\\
\Gamma_{b,b}\Gamma^{-1}_{b,a} &= \frac{-e^{g_t}}{e^{2g_t}+e^{g_t}-2}\frac{\sum_c e^{M_c+g_t 1_{c=a}}}{\sum_c e^{M_c+g_t 1_{c=b}}}\\
\Gamma_{b,d}\Gamma^{-1}_{d,a} &= \frac{-1}{e^{2g_t}+e^{g_t}-2}\frac{\sum_c e^{M_c+g_t 1_{c=a}}}{\sum_c e^{M_c+g_t 1_{c=b}}}\\
\Gamma_{a,a}\Gamma^{-1}_{a,a} &= \frac{(e^{g_t}+1)e^{g_t}}{e^{2g_t}+e^{g_t}-2}\\
2\Gamma_{a,d}\Gamma^{-1}_{d,a} &= \frac{-2}{e^{2g_t}+e^{g_t}-2}
\end{aligned}$$

Adding the right-hand sides of the first three equations yields zero and adding those of the last two gives one. This proves the formula for the inverse.

We now prove that the catastrophe manifold is the graph of χ. First, let us check that the range of χ is indeed the hyperplane H. Take an arbitrary point (m,β,t) in $H\times(0,\infty)\times(0,\infty)$ and let $\alpha=\chi(m,\beta,t)$.

$$\sum_{b=1}^{3}\alpha_b=\sum_{b=1}^{3}\sum_{a=1}^{3}m_a\Gamma^{-1}_{a,b}(\beta m,t) \tag{4.34}$$

Since $(1,1,1)^{\mathrm{T}}$ is an eigenvector of $\Gamma(\beta m,t)$ for the eigenvalue 1, it is also an eigenvector of $\Gamma^{-1}(\beta m,t)$ for the same eigenvalue. Therefore, we find

$$\sum_{b=1}^{3}\alpha_b=\sum_{a=1}^{3}m_a=1, \tag{4.35}$$

so α is an element of H. Next, we show that the catastrophe manifold is the graph of χ. The differential of $G_{\alpha,\beta,t}$ is given by

$$G'_{\alpha,\beta,t}(m)=\beta\left(m_a-\sum_{b=1}^{3}\alpha_b\Gamma_{b,a}(\beta m,t)\right)_{a=1}^{3} \tag{4.36}$$

Since $\Gamma(\beta m,t)$ is invertible, the equation $G'_{\alpha,\beta,t}(m)=0$ can be solved for α and we find $\alpha=\chi(m,\beta,t)$. Assume α is in Δ^2, then $G'_{\alpha,\beta,t}(m)=0$

implies that m also lies in Δ^2 since $0 < \Gamma_{b,a}(\beta m, t) < 1$ for all b, a in $\{1, 2, 3\}$.

To show (4.27) and (4.28) observe that the second derivative of $G_{\alpha,\beta,t}$ is given by the matrix

$$\begin{aligned} G''_{\alpha,\beta,t}(m) &= \beta \left(1_{a=b} - \beta \sum_{c=1}^{3} \alpha_c \frac{\partial \Gamma_{c,a}}{\partial M_b} \right)_{b,a=1}^{3} \\ &= \beta \left(1_{a=b} - \beta \sum_{c=1}^{3} \alpha_c \Gamma_{c,a} \Big(1_{a=b} - \Gamma_{c,b} \Big) \right)_{b,a=1}^{3} \end{aligned} \tag{4.37}$$

where $\Gamma = \Gamma(\beta m, t)$. The partial derivatives of $\Gamma_{c,a}$ are elements of the tangent space of Δ^2 for every c in $\{1, 2, 3\}$, that is, summing over a yields zero. Therefore:

$$\left\langle h, G''_{\alpha,\beta,t}(m) \frac{\partial}{\partial z} \right\rangle = \beta \sum_{b,a=1}^{3} 1_{a=b} h_b \left(\frac{\partial}{\partial z} \right)_a = \sum_a h_a \tag{4.38}$$

Since the coordinate basis of the (x, y, z)-chart is an orthogonal basis, we find $\langle \frac{\partial}{\partial x}, G''_{\alpha,\beta,t}(m) \frac{\partial}{\partial z} \rangle = \langle \frac{\partial}{\partial y}, G''_{\alpha,\beta,t}(m) \frac{\partial}{\partial z} \rangle = 0$ and $\langle \frac{\partial}{\partial z}, G''_{\alpha,\beta,t}(m) \frac{\partial}{\partial z} \rangle = \frac{3}{\beta}$. Since $\beta > 0$ and $\alpha = \chi(m, \beta, t)$, the condition for degenerate stationary points $\det G''_{\alpha,\beta,t}(m) = 0$ is equivalent to equation (4.28). □

4.4.2 Universality hypothesis connecting the mid-range dynamical model with the static model

In our work we are guided by the following *universality hypothesis*, which provides a useful organizing principle to understand the transitions which appear. It is suggested by the universality seen in local bifurcation theory, and verified for our model in the full set of mid-range temperatures $\beta < 3$, by means of our analytical treatment in the sequel of the paper, aided in some parts by computer algebra and numerics.

There exists a map from the two parameters temperature and time of the dynamical model to one effective temperature parameter of the static model of the form

$$(\beta, t) \mapsto \beta_{\mathrm{st}}(\beta, t) \tag{4.39}$$

which for our model is defined on the whole subset $\{(\beta,t)\,|\,0<\beta<3, t>0\}$ of the positive quadrant (and not only locally) and this map has the following property.

At fixed (β,t) the bifurcation set slice $B(\beta,t)\subset\Delta^2$, in the space of endconditionings α for the dynamical model, is diffeomorphic to a subset of the corresponding bifurcation set slice $B_{\mathrm{st}}(\beta_{\mathrm{st}})\subset\Delta^2$ of the static model under a smooth (β,t)-dependent map

$$\Delta^2\ni\alpha\mapsto\alpha_{\mathrm{st}}(\alpha,\beta,t). \tag{4.40}$$

See [38, Figure 2, page 973] for nine prototypical examples of such slices for the static model. Moreover the corresponding Maxwell sets of the dynamical and the static model get mapped onto each other by the same diffeomorphism. For corresponding values of (α,β,t) for the dynamical model and $(\beta_{\mathrm{st}},\alpha_{\mathrm{st}})$ the structure of stationary points of the rate functionals of the dynamical and the static model is identical. The image of Δ^2 under $\alpha_{\mathrm{st}}(\cdot,\beta,t)$, which we call the *effective observation window*, always contains the uniform distribution. However, it may be much smaller than Δ^2 for some parameter values. In fact, this will happen as $t\uparrow\infty$, as we will see. The map $\beta_{\mathrm{st}}(\beta,t)$ from dynamical to static parameters is (only) uniquely defined on the critical lines EW, B2B and BU of the dynamical model (see Figure 4.2) which get mapped to the corresponding static values $\beta_{\mathrm{st}}=4\log 2$, $\beta_{\mathrm{st}}=\frac{8}{3}$, and $\beta_{\mathrm{st}}=\frac{18}{7}$ [see 38, Table 1].

The following conjecture underlies this hypothesis, as it expresses the structural similarity of dynamical and static rate functional, by means of a parameter-dependent map acting on the state space Δ^2, compare with the definition of equivalent potentials in [49, Chapter 6, Section 1].

Conjecture 36. *There exists a set U which contains the unit simplex Δ^2 and is open in the hyperplane H such that*

(a) there exists a smooth map ψ_1 from the subset

$$D=\{(\alpha,\beta,t)\,|\,\beta<3, t>0, \alpha\in U\} \tag{4.41}$$

of the parameter space of the time-evolved model to the parameter space $(0,\infty)\times\Delta^2$ of the static model such that the map $(\alpha,\beta)\mapsto\psi_1(\alpha,\beta,t_0)$ is a diffeomorphism from $D\cap\{(\alpha,\beta,t):t=t_0\}$ to the respective image of this intersection for every $t_0>0$.

(b) there exists a smooth map ψ_2 *from* $D \times \Delta^2$ *to the state space* Δ^2 *of the static model such that the map* $m \mapsto \psi_2(\alpha, \beta, t, m)$ *defined on the interior of* Δ^2 *is a diffeomorphism onto its image for every* (α, β, t) *in* D.

(c) For every (α, β, t) *in* D *and every* m *in* Δ^2 *the following identity holds:*

$$G_{\alpha,\beta,t}(m) = f_{\psi_1(\alpha,\beta,t)} \circ \psi_2(\alpha, \beta, t, m) \tag{4.42}$$

where $f_{\beta,\alpha}$ *denotes the potential (5) of the static model [see 38, Sect. 1.2].*

(d) There exists a function $(\beta, t) \mapsto \beta_{\mathrm{st}}(\beta, t)$ *on* $(0, 3) \times (0, \infty)$ *such that*

$$\mathrm{pr}_1 \circ \psi_1(\alpha, \beta, t) = \beta_{\mathrm{st}}(\beta, t) \tag{4.43}$$

where pr_1 *denotes the projection* $(0, \infty) \times \Delta^2 \to (0, \infty)$. *In other words, the effective static inverse temperature* β_{st} *does not depend on the dynamical* α.

A comparison of Figure 4.9 with [38, Figure 5] gives evidence for the existence of the map ψ_1 as the bifurcation set slice of the static model looks structurally similar to the bifurcation set slice in a neighbourhood of the unit simplex of the dynamical model. The contour plots in the rightmost plots of the two figures support the existence of the map ψ_2 as the contour plot of the dynamical potential $G_{\alpha,\beta,t}$ looks structurally similar to a subset of the contour plot of the static potential $f_{\beta_{\mathrm{st}}(\beta,t),\alpha_{\mathrm{st}}(\alpha,\beta,t)}$. Note, however, that we are not going to construct the maps ψ_1 and ψ_2 in the following sections of the paper and we do not need to do it. Instead, we explicitly compute the critical lines from the dynamical potential following the ideas of singularity theory. This means that the lines can be found independently of the construction of the maps ψ_1 and ψ_2. The behavior of the model in the vicinity of these lines follows from Thom's classification theorem [see 45, Section 5 of Chapter 3] and our global analysis is supported by the global numerical analysis of the relevant parts of the dynamical bifurcation set. In the following sections we now proceed with the discussion of the critical lines.

4.4.3 The symmetric cusp exit (SCE) line and the non-Gibbs temperature

The non-Gibbs inverse temperature β_{NG} is defined as the supremum of all β such that $\mu_{n,\beta,t}$ is sequentially Gibbsian for all positive t. It turns out to be a maximum. As the type of transitions of the dynamical model for mid-range temperatures can be understood in terms of the static case, let us remark that in the static Potts model the first type of bad magnetic fields that show up with increasing β are due to three symmetric cusp singularities, the "rockets" [see 38, Figures 2 and 4] and that there are no bad magnetic fields for $\beta \leq 2$. Therefore, in the dynamical model, we look for symmetric cusp points that have just passed the simplex edges in their midpoint and moved outside, which leads us to the *symmetric cusp exit line* in the dynamical phase diagram. Without loss of generality by symmetry we consider the simplex edge where $\alpha_1 = 0$.

Proposition 37. *Fix any positive β and t, let m be a point in H with (x, y, z)-coordinates $(0, y, 0)$.*

(a) The point $\alpha = \chi(m, \beta, t)$ in H is a symmetric cusp point on the simplex edge if and only if

$$\frac{6}{\beta}y + \frac{e^{g_t} + 1 - 2e^{3y}}{e^{g_t} + 1 + e^{3y}} = 0 \tag{4.44}$$

$$\frac{6}{\beta} + \frac{3(e^{g_t} - 1)^2}{(e^{g_t} + 1 + e^{3y})^2} - \frac{3(e^{g_t} + 1)}{e^{g_t} + 1 + e^{3y}} = 0. \tag{4.45}$$

(b) The solutions of the system (4.44–4.45) can be explicitly parametrized in the form

$$\beta = \frac{2s(2e^s + F(s))}{4e^s - F(s)} \tag{4.46}$$

$$g_t = \log\left(\frac{1}{2}F(s) - 1\right) \tag{4.47}$$

where

$$F(s) = -(s-1)e^s - 4s + \sqrt{((s-1)e^s + 4s)^2 + 8(2s + e^{2s})}. \tag{4.48}$$

for $s < 0$.

(c) The non-Gibbs temperature is given via

$$\beta_{\mathrm{NG}} = \frac{2s_0(2e^{s_0} + F(s_0))}{4e^{s_0} - F(s_0)} \approx 2.52885 \tag{4.49}$$

where s_0 is the unique zero in $(-\infty, 0)$ of

$$\frac{64\,s^3 + 64\,s^2 + s(s^2+s+6)e^{3\,s} + 4\,s(5\,s+6)e^{2\,s} - 8\,s(2\,s-3)e^{s}}{\sqrt{((s-1)e^s + 4s)^2 + 8(2s + e^{2s})}} \\ -16\,s^2 - s(s+2)e^{2\,s} + 4\,s(s-2)e^{s} - 8\,s. \tag{4.50}$$

Proof. Let us first prove item 1. A symmetric cusp point α is the image of a symmetric degenerate stationary point m under the map $\chi(\cdot, \beta, t)$ at which the tangent vector of the curve of degenerate stationary points (given by vanishing Hessian determinant) is parallel to the direction of degeneracy. The partial derivatives of $G_{\alpha,\beta,t}$ with respect to x and z vanish at m because of symmetry, so it is sufficient for a stationary point m to have a vanishing partial derivative with respect to the y-coordinate of m. Now, for the gradient we note that

$$\begin{aligned} \frac{\partial G_{\alpha,\beta,t}}{\partial y} &= 2m_1 - m_2 - m_3 - \sum_{b=1}^{3} \alpha_b (2\Gamma_{b,1} - \Gamma_{b,2} - \Gamma_{b,3}) \\ &= \frac{6}{\beta} y - \sum_{b=1}^{3} \alpha_b (3\Gamma_{b,1} - 1) \\ &= \frac{6}{\beta} y + 1 - \frac{3e^{3y}}{e^{3y} + e^{g_t} + 1} \end{aligned} \tag{4.51}$$

where we have abbreviated $\Gamma_{b,a} = \Gamma_{b,a}(\beta m, t)$ and used the fact that α lies on the simplex edge $\alpha = (0, \frac{1}{2}, \frac{1}{2})$. This yields Equation (4.44).

We will now derive equation (4.45). Note that the mixed partial derivative, which appears in the degeneracy condition (4.28), vanishes at partially symmetric points:

$$\frac{\partial^2 G_{\alpha,\beta,t}}{\partial x \partial y} = -3 \sum_{b=1}^{3} \alpha_b \frac{\partial \Gamma_{b,1}}{\partial x} = 3\sqrt{3} \sum_{b=1}^{3} \alpha_b \Gamma_{b,1} (\Gamma_{b,3} - \Gamma_{b,2}) \tag{4.52}$$

Plugging in $\alpha = (0, \frac{1}{2}, \frac{1}{2})$, the right-hand side of the last equality in (4.52) vanishes because $\Gamma_{3,3} - \Gamma_{3,2} = \Gamma_{2,2} - \Gamma_{2,3}$ for points m which have the partial symmetry $m_2 = m_3$. Therefore the degeneracy condition (4.28) is in product form. We calculate the remaining partial derivatives:

$$\frac{\partial^2 G_{\alpha,\beta,t}}{\partial y^2} = \frac{6}{\beta} - 9(\Gamma_{2,1} - \Gamma_{2,1}^2) = 9\left(\Gamma_{2,1} - \frac{1}{2}\right)^2 - \frac{9}{4} + \frac{6}{\beta} \tag{4.53}$$

$$\frac{\partial^2 G_{\alpha,\beta,t}}{\partial x^2} = \frac{6}{\beta} - 3\Big(\Gamma_{2,2} + \Gamma_{2,3} - (\Gamma_{2,3} - \Gamma_{2,2})^2\Big) \tag{4.54}$$

The partial derivative (4.53) is always positive for $\beta < \frac{8}{3}$. This means we only have to consider the zeros of (4.54). This yields equation (4.45).

We will now explain the parametrization of the set of solutions given in 2. First note that the variable β can be eliminated from Equation (4.45) using Equation (4.44) for all $y \neq 0$. When we set $w = e^{g_t} + 1$ we find that the resulting equation is a quotient of quadratic polynomials in w:

$$-\frac{w^2 + ((3y-1)e^{3y} + 12y)w - 2(6y + e^{6y})}{y(w + e^{3y})^2} = 0 \tag{4.55}$$

Since $w > 2$, it suffices to consider the numerator of the left-hand side. The discriminant of this quadratic polynomial is given by

$$D = ((3y-1)e^{3y} + 12y)^2 + 8(6y + e^{6y}). \tag{4.56}$$

It is positive for all real y. Therefore, this polynomial has two real roots. Because $w > 2$, we choose the larger of the two solutions

$$w = \frac{1}{2}\left(-(3y-1)e^{3y} - 12y + \sqrt{D}\right) = \frac{1}{2}F(s) \tag{4.57}$$

where we have defined $s = 3y$ and used the definition of $F(s)$ in Equation (4.48). Furthermore, $F(s) > 4$ for $s \neq 0$ such that Equation (4.47) yields positive values for g_t.

Finally, the non-Gibbs inverse temperature is the minimal value of β along the curve given by the parametrization (4.46–4.47). Therefore we calculate the derivative of (4.46) which gives

$$\frac{\mathrm{d}\beta}{\mathrm{d}s} = -2 \cdot \frac{2(3s-1)e^s F(s) - 6se^s F'(s) + F^2(s) - 8e^{2s}}{(4e^s - F(s))^2}. \tag{4.58}$$

Since $4e^s - F(s)$ is never zero for any s in $(-\infty, 0)$, we only have to consider the numerator of the fraction. We calculate the derivative of F

$$F'(s) = -se^s - 4 + \frac{((s-1)e^s + 4s)(4 + se^s) + 8(1 + e^{2s})}{\sqrt{((s-1)e^s + 4s)^2 + 8(2s + e^{2s})}}. \tag{4.59}$$

Plugging everything together, $\mathrm{d}\beta/\mathrm{d}s = 0$ is exactly fulfilled for the zero of the function defined in (4.50). □

Lemma 38. *Suppose β lies in the interval $(\beta_{\mathrm{NG}}, 3)$. The entry time $t_{\mathrm{NG}}(\beta)$ into the non-Gibbs region is given by*

$$t_{\mathrm{NG}}(\beta) = \frac{1}{3} \log \frac{2(\beta - 3y)e^{3y} + \beta + 6y}{2((\beta - 3y)e^{3y} - \beta - 6y)} \tag{4.60}$$

where y is the largest root in $(-\frac{\beta}{6}, 0)$ of

$$\begin{aligned} y \mapsto & 2\,\beta^2 + 24\,\beta y + 72\,y^2 - \left(\beta^2 + 3\,\beta y - 18\,y^2 - 9\,\beta\right) e^{6\,y} \\ & - 4\left(\beta^2 + 3\,\beta y - 18\,y^2\right) e^{3\,y}. \end{aligned} \tag{4.61}$$

Proof. The entry time t_{NG} is given by the first entry of *rockets* into the unit simplex while increasing the time t and keeping β fixed. This is because, if the pentagrams unfold at all under increase of time, they unfold after the rockets have entered the unit simplex Δ^2. This will be clear in the next subsection where we compute the butterfly line. So let us consider the system (4.44–4.45) and fix any positive $\beta < 3$. Since the relation (4.4) between g_t and t is strictly monotonically decreasing, we have to look for the maximal g_t such that (β, g_t, y) with negative y is a solution to the system (4.44–4.45), which defines the symmetric cusp exit line. Here, y is a magnetization-type variable. We can solve Equation (4.44) for $w = e^{g_t} + 1$ to obtain

$$w = \frac{2(\beta - 3y)e^{3y}}{\beta + 6y}. \tag{4.62}$$

Plugging this into the left-hand side of the degeneracy condition (4.45), we arrive at

$$\begin{aligned} \frac{2e^{-6y}}{3\beta^2}\Big(2\beta^2 + 24\beta y + 72y^2 - (\beta^2 + 3\beta y - 18y^2 - 9\beta)e^{6y} \\ - 4(\beta^2 + 3\beta y - 18y^2)e^{3y}\Big) = 0. \end{aligned} \tag{4.63}$$

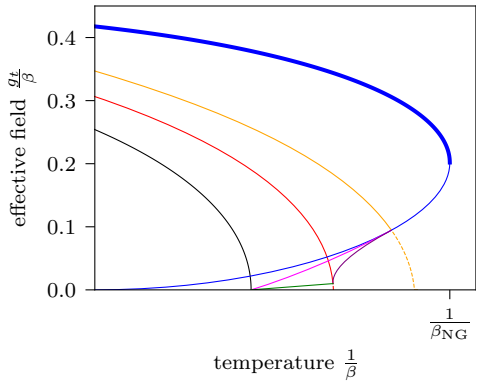

Figure 4.5: The thick blue line, which ends at the non-Gibbs temperature $\frac{1}{\beta_{\mathrm{NG}}}$, marks the entry time in the dynamical phase diagram. Time is a monotonically decreasing function of g_t so the first time we hit the symmetric cusp exit line when moving on a vertical line of fixed temperature corresponds to the entry time.

This yields the expression of (4.61). Since the right-hand side of (4.62) is increasing with y, we have to pick the largest root of (4.61). □

4.4.4 The butterfly unfolding (BU) line and butterfly exit temperature

The unfolding of the pentagrams is a very important mechanism since it changes the set of bad empirical measures from straight lines to Y-shaped, branching curves. This mechanism is already present in the static case, however, in contrast to the static case we have to deal with the fact that in some parameter regions the pentagrams do not fully lie inside of the unit simplex. This leads us to the definition of a *butterfly exit inverse temperature* β_{BE} for which at some point in time $t > 0$ there is a cusp point on an edge of the simplex that is about to unfold into a pentagram. By definition, β_{BE} lies between β_{NG} and $\frac{8}{3}$. The value $\frac{8}{3}$ is the first inverse temperature for which a beak-to-beak scenario inside of the unit simplex appears as we will see in Section 4.4.7.

Proposition 39. *Let $v(m, \beta, t) = (\varphi_\beta)_2 \circ \chi(m, \beta, t)$ be the parallel coordinate of $\chi(m, \beta, t)$ and let $\beta(s)$ and $t(s)$ be given by* (4.46–4.47)*. The*

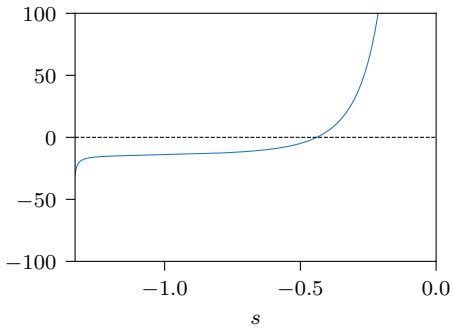

Figure 4.6: This figure shows a plot of the function (4.65) which is involved in the expression for the *butterfly exit (BE) temperature* in Proposition 39.

butterfly exit β_{BE} *is given by*

$$\beta_{\mathrm{BE}} = \frac{2s_0(2e^{s_0} + F(s_0))}{4e^{s_0} - F(s_0)} \approx 2.59590 \tag{4.64}$$

where $s_0 < 0$ *is the largest zero of*

$$s \mapsto \frac{\partial^2 v}{\partial x^2}\Big(m(s), \beta(s), t(s)\Big) + \frac{\partial v}{\partial y}\Big(m(s), \beta(s), t(s)\Big)\ddot{\gamma}_s(0) \tag{4.65}$$

and γ_s *is the implicit function* $y = \gamma_s(x)$ *defined in a neighbourhood of* $(x, y) = (0, \frac{s}{3})$ *by the degeneracy condition* (4.28)*.*

Note that equation (4.65) is explicitly computed by a computer program because its expression is very complicated. Nevertheless it is possible to plot the function (see Figure 4.6).

Proof. Let us first fix β between β_{NG} and $\frac{8}{3}$ and a positive t. Consider a point α on the midpoint of one of the edges of Δ^2 such that (α, β, t) belongs to the bifurcation set. Furthermore, without loss of generality by symmetry let us assume that $\alpha_2 = \alpha_3$. To this point corresponds a degenerate stationary point m that has the same symmetry $m_2 = m_3$. We can solve the degeneracy condition (4.28) in a neighbourhood of m in the form $y = \gamma_{\beta,t}(x)$ such that $\gamma_{\beta,t}(0)$ is the y-coordinate of m. In α-space in a neighbourhood of $\alpha = \chi(m, \beta, t)$ we can now write the

bifurcation set as $\chi(\varphi_\beta^{-1}(x, \gamma_{\beta,t}(x), 0), \beta, t)$. We know that the parallel component v of α fulfills

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} v(\gamma_{s_0}(x), \beta_*, t_*) = 0 \tag{4.66}$$

when we follow the curve γ_s through the bifurcation set. This is because it has a minimum before the pentagram unfolds and it has a maximum after the pentagram has unfolded. The curve γ of degenerate stationary points is obtained by solving equation (4.28) in the form $y = \gamma(x)$ around $(0, y^*)$ where y^* is the parallel component of m^*. Let us now compute the second derivative of the v-component of the curve:

$$\begin{aligned} \frac{\mathrm{d}^2 v}{\mathrm{d}x^2}\Big|_{x=0} &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial v}{\partial x} + \frac{\partial v}{\partial y}\dot{\gamma}(x)\right) \\ &= \frac{\partial^2 v}{\partial x^2}\Big|_{x=0} + \frac{\partial v}{\partial y}\Big|_{x=0} \ddot{\gamma}(0) \end{aligned} \tag{4.67}$$

The other mixed partial derivatives of v vanish since $\dot{\gamma}(0) = 0$ because of symmetry. □

Furthermore, we compute $\ddot{\gamma}(0)$ via implicit differentiation: Let us write $f(x, y)$ for the left-hand side of (4.28) viewed as a function in the unit simplex in (x, y)-coordinates. By implicit differentiation we then find:

$$\dot{\gamma}(x) = -\frac{\partial f}{\partial x} \Big/ \frac{\partial f}{\partial y} \tag{4.68}$$

And therefore:

$$\begin{aligned} \ddot{\gamma}(0) &= -\frac{\frac{\partial^2 f}{\partial x^2}}{\frac{\partial f}{\partial y}} + \frac{\frac{\partial f}{\partial x}\frac{\partial^2 f}{\partial x \partial y}}{\left(\frac{\partial f}{\partial y}\right)^2} = -\frac{\frac{\partial^2 f}{\partial x^2}}{\frac{\partial f}{\partial y}} - \dot{\gamma}(0)\frac{\frac{\partial^2 f}{\partial x \partial y}}{\frac{\partial f}{\partial y}} \\ &= -\frac{\partial^2 f}{\partial x^2} \Big/ \frac{\partial f}{\partial y}. \end{aligned} \tag{4.69}$$

Using the symbolic calculus tools (see page 107) we can obtain an expression for (4.65).

Using a similar approach it is possible to compute the line in the dynamical phase diagram for which we find butterfly points no matter

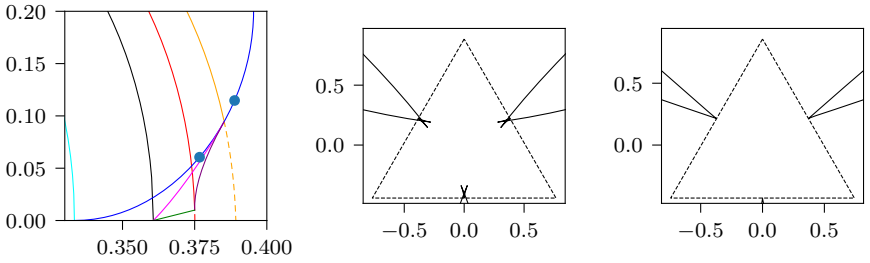

Figure 4.7: This figure shows the bifurcation set sliced at two points on the symmetric cusp exit line. The left plot shows a slice before the butterfly exit point is passed (lower point in the phase diagram), the right plot shows a slice after the butterfly exit point (intersection point of yellow and blue line) on the symmetric cusp exit line.

where these points are with respect to the unit simplex. The key idea that the parallel component of the curve in α-space has a vanishing second derivative with respect to the curve parameter stays the same. But since we do not restrict the point in α-space to lie on the unit simplex we lose one equation and we end up with a one-dimensional set of solutions.

Proposition 40. *For β in $(\beta_{\mathrm{BE}}, \frac{8}{3})$ the butterfly unfolding happens at the unique* butterfly transition time *$t_{\mathrm{BU}}(\beta)$ which is obtained as follows: Define a function H via*

$$H(\beta, s) = H_1(\beta, s) + \sqrt{H_2(\beta, s)} \tag{4.70}$$

where

$$\begin{aligned}
H_1(\beta, s) &= \beta e^{2s} - s e^{2s} + 4\beta e^{s} - 4s e^{s} + \beta + 2s - 3e^{2s} - 3e^{s} \\
H_2(\beta, s) &= \Big(\beta^2 - 2\,(\beta - 3)s + s^2 - 6\,\beta + 9\Big)e^{4\,s} \\
&\quad + 2\,\Big(4\,\beta^2 - (8\,\beta - 9)s + 4\,s^2 - 9\,\beta - 9\Big)e^{3\,s} \\
&\quad + 3\,\Big(6\,\beta^2 - 2\,(5\,\beta - 6)s + 4\,s^2 - 18\,\beta + 3\Big)e^{2\,s} \\
&\quad + 2\,\Big(4\,\beta^2 + 2\,(2\,\beta - 15)s - 8\,s^2 - 15\,\beta\Big)e^{s} \\
&\quad + \beta^2 + 4\,\beta s + 4\,s^2
\end{aligned}$$

and a function

$$t(\beta, s) = \frac{1}{3} \log \frac{H(\beta, s) + 6e^s}{H(\beta, s) - 12e^s}. \tag{4.71}$$

Then the butterfly transition time $t_{\mathrm{BU}}(\beta)$ *is given by*

$$t_{\mathrm{BU}}(\beta) = t(\beta, s_*(\beta)) = \frac{1}{3} \log \frac{H(\beta, s_*(\beta)) + 6e^{s_*(\beta)}}{H(\beta, s_*(\beta)) - 12e^{s_*(\beta)}} \tag{4.72}$$

and $s_*(\beta) < 0$ *is the largest zero of*

$$s \mapsto \frac{\partial^2 v}{\partial x^2}\Big(\varphi_\beta(0, \frac{s}{3}, 0), \beta, t(\beta, s)\Big) + \frac{\partial v}{\partial y}\Big(\varphi_\beta(0, \frac{s}{3}, 0), \beta, t(\beta, s)\Big)\ddot{\gamma}_{\beta, t(\beta,s)}(0). \tag{4.73}$$

Proof. Using the same reasoning as in the proof of Proposition 39, we find that the point m maps under $\chi(\cdot, \beta, t)$ to a point α that is about to unfold into a pentagram if

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg|_{x=0} v(\varphi_\beta^{-1}(x, \gamma_{\beta,t}(x), 0), \beta, t) = 0 \tag{4.74}$$

where $\gamma_{\beta,t}$ is obtained by solving the degeneracy condition (4.28) in the form $y = \gamma_{\beta,t}(x)$ in a neighbourhood of the point m. This equation is now dependent on m, β and t, that is, we have one equation and three variables (m is one-dimensional because $m_2 = m_3$). Additionally, since we know that the direction of degeneracy is the x-direction, we have the equation

$$\frac{\partial^2 G_{\alpha,\beta,t}}{\partial x^2}\bigg|_{x=0} = 0. \tag{4.75}$$

This equation can be solved for $w = e^{g_t} + 1$ which yields (4.70). Plugging this into (4.74), we are left to find the zeros of (4.73) for some fixed β in the interval $(\beta_{\mathrm{BE}}, \frac{8}{3})$. □

4.4.5 Reentry into Gibbs: the asymmetric cusp exit (ACE) line

In the β-regime $(\beta_{\mathrm{NG}}, \beta_{\mathrm{BE}})$, three pentagrams unfold inside of the simplex at an intermediate time and leave the simplex as t increases further.

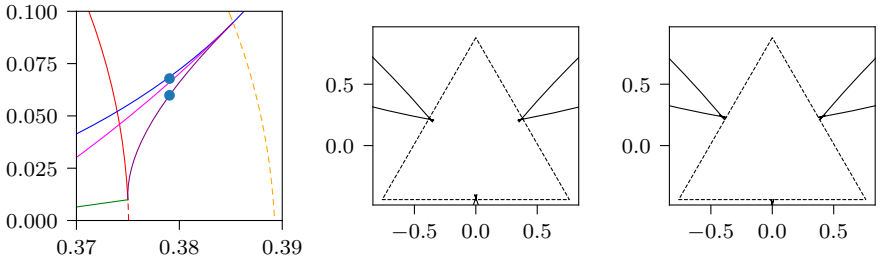

Figure 4.8: This figure shows two bifurcation set slices that illustrate the exit of the asymmetric cusp points. The central plot shows the bifurcation set slice for a time at which the exit has not yet happened (upper point in the phase diagram). The rightmost plot shows the bifurcation set slice exactly on the purple line ACE, that is, when the exit is just happening.

Since we are interested in phase-coexistence of the first layer model $\bar{\mu}_n$ (Lemma 33) and the phase-coexistence lines of the pentagram end in the asymmetric cusp points of the pentagrams, we must compute the exit time $t_{\mathrm{G}}(\beta)$ of these points for β in the above regime. Like in the previous subsection, this is done using a combination of symbolic and numerical computation (see page 107). First, let us state the problem that we need to solve.

Proposition 41. *Fix a positive β and positive t and consider the set of solutions m to the degeneracy condition* (4.28) *with $\alpha = \chi(m, \beta, t)$.*

(a) There is exactly one branch of solution with $m_2 = m_3$ and it is given by the graph of a map $x \mapsto y = \gamma_{\beta,t}(x)$.

(b) Furthermore, define the map $(x, y) \mapsto v(x, y)$ via

$$v(x, y) = (\varphi_\beta)_1 \circ \chi(\varphi_\beta^{-1}(x, y, 0), \beta, t). \tag{4.76}$$

Then the asymmetric cusps of the pentagrams are on the simplex edges if and only if

$$v(x, \gamma_{\beta,t}(x)) = -\frac{1}{6}\beta \tag{4.77}$$

$$\frac{\partial v}{\partial x}(\varphi_\beta(x, \gamma_{\beta,t}(x), 0)) + \frac{\partial v}{\partial y}(\varphi_\beta(x, \gamma_{\beta,t}(x), 0))\dot{\gamma}_{\beta,t}(x) = 0. \tag{4.78}$$

Proof. The location of the asymmetric cusps of the pentagrams on the curve $x \mapsto \chi(\varphi_\beta^{-1}(x, \gamma_{\beta,t}, 0), \beta, t)$ are given by the local maxima of the parallel component $v(x)$ as a function of the curve parameter x (see Figure 4.8). This yields (4.78). Equation (4.77) comes from the constraint that the cusp point lies on the simplex edge because for points on the edge the parallel component equals $-\frac{1}{6}\beta$ in the chart (4.26). □

Now, similarly to the case for the butterfly line, the computation of $\dot{\gamma}_{\beta,t}(x)$ by hand is impractical. Therefore we compute the expression symbolically with the help of the computer. This allows us to numerically determine the course of the line in the dynamical phase diagram. Now, because it is impossible to solve the degeneracy equation (4.28) in the form $y = \gamma_{\beta,t}(x)$ explicitly, we proceed as follows. Note that it is possible to solve (4.77) for β and plug it into equation (4.78). We then fix some value of g_t, and numerically solve the system consisting of the degeneracy condition (4.28), where β is substituted from (4.77), and equation (4.78), where $\gamma_{\beta,t}$ is substituted by y and

$$\dot{\gamma}_{\beta,t}(x) = -\frac{\partial f}{\partial x} \Big/ \frac{\partial f}{\partial y} \tag{4.79}$$

where f denotes the left-hand side of (4.28) considered as a function of (x, y). This yields two equations in the two variables x and y.

4.4.6 The triple point exit (TPE) line

To each of the three pentagrams there belongs a special point, the *triple point* [see 38, Sects. 3.2]. This point is characterized by the coexistence of three global minima, that is, the functional values of all the three minimizers are equal. First, we discuss the existence of these points and then we determine for each fixed positive β the exit time $t_{\text{triple}}(\beta)$. This is the last time for which there are bad empirical measures with partial symmetry that lie inside the unit simplex.

Proposition 42. *For each pair* (β, t) *in*

$$\{(\beta, t) \mid \beta_{\text{BE}} < \beta < 4 \log 2, t > t_{\text{BU}}(\beta)\} \tag{4.80}$$

there exists exactly one α *in the hyperplane* H *with* $\alpha_1 \leq \alpha_2 \leq \alpha_3$ *such that* $G_{\alpha,\beta,t}$ *has precisely three global minimizers.*

Proof. By symmetry, the triple point α has the partial symmetry $\alpha_2 = \alpha_3$. Therefore consider the curve $v \mapsto \alpha(v) = \varphi_\beta^{-1}(0, v, 0)$ which crosses the α-region for which the potential $G_{\alpha,\beta,t}$ has three minimizers two of which lie inside the same fundamental cell $m_1 \leq m_2 \leq m_3$. There is always such a region because the pentagrams have already unfolded ($t > t_{\text{but}}$). This gives rise to the two maps $v \mapsto m(v)$ and $v \mapsto m'(v)$ which map v to one of the two minimizers $m(v)$ or $m'(v)$ inside this cell. Assume that $\varphi_\beta(m(v)) = (x(v), y(v), 0)$ and $\varphi_\beta(m'(v)) = (0, y'(v), 0)$ with $y'(v) > y(v)$ and $x(v) > 0$. Now, we can define the difference

$$g(v) := G_{\alpha(v),\beta,t}(m(v)) - G_{\alpha(v),\beta,t}(m'(v)) \tag{4.81}$$

for all v such that $\alpha(v)$ lies in the former regime. Therefore

$$\begin{aligned} g'(v) &= \frac{\partial G_{\alpha(v),\beta,t}}{\partial v} - \frac{\partial G_{\alpha(v),\beta,t}}{\partial v} \\ &= \log \frac{(e^{g_t+2x} + e^{3y+x} + 1)(e^{g_t+3y'} + 2)^2(e^{g_t} + e^{2x} + e^{3y+x})}{(e^{g_t+x+3y} + e^{2x} + 1)^2(e^{g_t} + e^{3y'} + 1)^2} \end{aligned} \tag{4.82}$$

since $m(v)$ and $m'(v)$ are stationary points. □

Since the pentagrams in the bifurcation slices leave the simplex (observation window), it is necessary for a discussion of the bad empirical measures that we find the time when the triple points leave the unit simplex. The problem that we have to solve is stated in the following proposition.

Proposition 43. *Fix any positive β in the interval $(\beta_{\text{BE}}, 4\log 2)$ and let α be the midpoint of the edge of the simplex with $\alpha_2 = \alpha_3$. First, define the function*

$$t(\beta, y) = \frac{1}{3}\log\left(\frac{2(\beta - 3y)e^{3y} + \beta + 6y}{2((\beta - 3y)e^{3y} - \beta - 6y}\right) \tag{4.83}$$

The exit time $t_{\text{TPE}}(\beta)$ is then given by $t_{\text{TPE}}(\beta) = t(\beta, y'(\beta))$ where $\varphi_\beta(0, y'(\beta))$ and $\varphi_\beta(x(\beta), y(\beta))$ lie in the fundamental cell $m_1 \leq m_2 \leq$

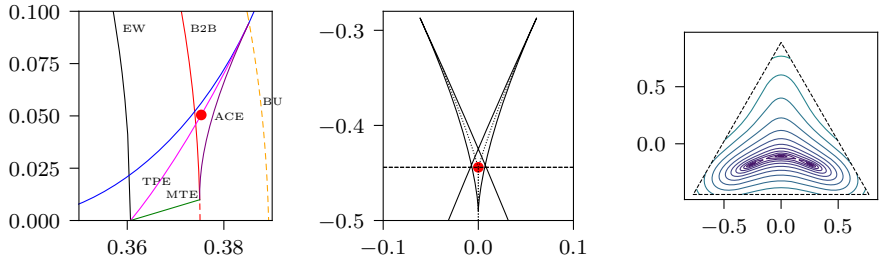

Figure 4.9: The four-dimensional parameter (α, β, t) is represented by the two red dots in the two plots on the left. The first of these plots displays a region of the dynamical phase diagram and the second plot the respective bifurcation set slice clipped to a rectangle near the lower simplex edge which is represented by the dashed horizontal line. The rightmost plot shows contour lines of the potential $G_{\alpha,\beta,t}$ for the respective parameter. As expected for a triple point, the contour lines show three equally deep minimizers of the potential.

m_3 and the triple $(y'(\beta), x(\beta), y(\beta))$ is a solution to the following system of equations.

$$G_{\alpha,\beta,t(\beta,y')} \circ \varphi_\beta(0, y', 0) = G_{\alpha,\beta,t(\beta,y')} \circ \varphi_\beta(x, y, 0) \quad (4.84)$$

$$(\varphi_\beta)_1 \circ \chi((\varphi_\beta)^{-1}(0, y', 0), \beta, t(\beta, y')) = (\varphi_\beta)_1 \circ \chi((\varphi_\beta)^{-1}(x, y, 0), \beta, t(\beta, y')) \quad (4.85)$$

$$(\varphi_\beta)_2 \circ \chi((\varphi_\beta)^{-1}(0, y', 0), \beta, t(\beta, y')) = (\varphi_\beta)_2 \circ \chi((\varphi_\beta)^{-1}(x, y, 0), \beta, t(\beta, y')) \quad (4.86)$$

Note that the expressions of the equations (4.84 – 4.86) are computed symbolically by the computer (see page 107 for more information). They are not displayed here because of their length. Figure 4.9 shows a contour plot of the HS transform $G_{\alpha,\beta,t}$ with $\alpha = (0, \frac{1}{2}, \frac{1}{2})$ and (β, t) on the line TPE.

Proof. The system of equations mainly comes from two ingredients: equal depth of two minimizers and same end-conditioning α for these two minimizers. The triple point is characterized by a coexistence of three global minimizers and since a triple point α must fulfill the symmetry relation $\alpha_2 = \alpha_3$, we find that it is sufficient to compare the two minimizers in the fundamental cell $m_1 \leq m_2 \leq m_3$. Because

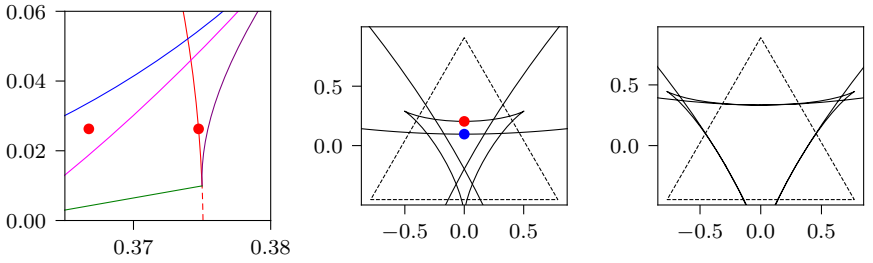

Figure 4.10: The beak-to-beak mechanism is characterized by the merging of two horns of two different pentagrams. This merging joins two connected components of the complement of the bifurcation set slice when crossing the red line from right to left. As can be seen in the two rightmost plots, this merging happens on the axis of symmetry. The red dots in the dynamical phase diagram on the left mark the time–temperature pairs that correspond to the bifurcation set slices from left to right. The dots in the central plot correspond to the points of the same color in Figure 4.11.

$\alpha_2 = \alpha_3$, we always have one symmetric stationary point so that the two minimizers have the coordinates $(0, y', 0)$ and $(x, y, 0)$. Since we now that either minimizer is a stationary point, we can use the vanishing of the first partial derivative of $G_{\alpha,\beta,t}$ with respect to the y-coordinate to eliminate the time variable t from the equations. This yields the function in equation (4.83). Using this function we can eliminate the variable t from the equal depth condition and the other two equations that require that the minimizers belong to the same end-conditioning α. □

4.4.7 The beak-to-beak (B2B) line

The beak-to-beak point in the static model is characterized as a cusp point that lies in a segment from the center of the simplex to one vertex, that is, for example it has $y > 0$. The following proposition describes the line of beak-to-beak points and a parametric representation in terms of roots of a cubic polynomial. Note that, despite the fact that the line continues to exist for $\beta > 3$, the structural behavior of the bifurcation set around the beak-to-beak point might change in the regime $\beta > 3$.

Proposition 44. *Fix any positive β and t, let m be a point in H with coordinates $(0, y, 0)$.*

(a) The point $\alpha = \chi(m, \beta, t)$ is a beak-to-beak point if and only if

$$-(\beta + 6y - 2)(e^{g_t} + 1)e^{-3y} - (\beta - 3y - 1)e^{g_t+3y} + e^{g_t}(e^{g_t} + 1) = 0 \tag{4.87}$$

$$(\beta + 6y - 4)(e^{g_t} + 1)e^{-3y} - (\beta - 3y - 2)e^{g_t+3y} = 0 \tag{4.88}$$

(b) The solutions to this system can be parametrized in terms of $s = 3y$ in the form

$$\beta = \frac{2(s-2)w_*(s) + (s+2)(w_*(s) - 1)e^{2s}}{(w_*(s) - 1)e^{2s} - w_*(s)} \tag{4.89}$$

$$g_t = \log(w_*(s) - 1) \tag{4.90}$$

where $s > s_* \approx 0.66656$ and $w_*(s)$ is the unique root in the interval $(2, \infty)$ of the cubic polynomial

$$\begin{aligned}(e^{3s} - e^s)\, w^3 - (6se^{2s} + e^{4s} + 2e^{3s} - 3e^{2s} - e^s - 2)\, w^2 + \\ (6se^{2s} + 2e^{4s} + 3e^{3s} - 3e^{2s} - 2e^s)\, w - e^{4s} - 2e^{3s}.\end{aligned} \tag{4.91}$$

The positive real number s_* is the unique root in $(0, \infty)$ of the function

$$s \mapsto -12se^{2s} - e^{4s} + 4e^{3s} + 6e^{2s} - 8e^s + 8. \tag{4.92}$$

(c) The beak-to-beak point enters the simplex for $s = 2/3 > s_*$ at which $\beta = \frac{8}{3}$ and $g_t \approx 0.026481$.

Proof. From the analysis of the static model [see 38, Figure 2, rightmost plot of the first row and neighbouring plots for smaller or larger β] we know that the beak-to-beak point (α_*, β_*, t_*) is such that if we fix $\alpha = \alpha_*$ but change the parameters β or t we either find that $\alpha = \alpha_*$ is contained in a cell with two minimizers or in a cell with one minimizer. Since α_* lies on the axis of symmetry, we know $\alpha_* = \chi(m_*, \beta_*, t_*)$ where m_* lies on the axis of symmetry as well, and we find in coordinates $\varphi_\beta(\alpha_*) = (0, v(m_*, \beta_*, t_*), 0)$, so it suffices to study

$$v(m, \beta, t) = (\varphi_\beta)_2 \circ \chi(m, \beta, t) = \frac{(\beta + 6y)we^{-3y} - (\beta - 3y)(w - 1)e^{3y} + 3(w^2 - w + 2)y - \beta}{3(w^2 - w - 2)} \tag{4.93}$$

as a function of the y-coordinate of m. As before substitute $w = e^{g_t} + 1$. In Figure 4.11 you see a minimum and a maximum collide and form a saddle point. This is exactly the beak-to-beak behavior. The point (β, t) for which this collision has just happened is given by the vanishing of the first and second derivatives of $v(m, \beta, t)$ with respect to the y-coordinate of m. Now, the derivatives are given by:

$$\frac{\mathrm{d}v}{\mathrm{d}y}(m,\beta,t) = \frac{-(\beta+6y-2)we^{-3y} - (\beta-3y-1)(w-1)e^{3y} + w^2 - w + 2}{w^2 - w - 2} \tag{4.94}$$

$$\frac{\mathrm{d}^2v}{\mathrm{d}y^2}(m,\beta,t) = \frac{3(\beta+6y-4)we^{-3y} - 3(\beta-3y-2)(w-1)e^{3y}}{w^2 - w - 2} \tag{4.95}$$

Since $w > 2$, it suffices to consider the numerators of the above expressions. This yields equations (4.87) and (4.88).

Let us now prove the parametric form of the solutions. Equation (4.88) is linear in β as long as $e^{g_t} + 1 - e^{g_t+6y} \neq 0$ and can then be solved for β to yield (4.89) after substituting $w = e^{g_t} + 1$ and $s = 3y$. Suppose now $e^{g_t} + 1 - e^{g_t+6y} = 0$ which is equivalent to $e^{g_t} = \frac{1}{e^{6y}-1}$. Equation (4.88) would in this case read

$$\frac{(9y-2)e^{3y}}{e^{6y}-1} = 0 \tag{4.96}$$

which is only fulfilled for $y = \frac{2}{9}$. However, this leads to the contradiction $e^{g_t} = \frac{1}{e^{\frac{4}{3}}-1} < 1$ but $g_t > 0$. Therefore, we can assume that we can solve (4.88) for β. Plugging this into equation (4.87) we arrive at the following fraction of polynomials in w.

$$\frac{\begin{aligned}&\left(e^{3s} - e^{s}\right)w^3 - \left(6se^{2s} + e^{4s} + 2e^{3s} - 3e^{2s} - e^{s} - 2\right)w^2 \\ &\qquad + \left(6se^{2s} + 2e^{4s} + 3e^{3s} - 3e^{2s} - 2e^{s}\right)w - e^{4s} - 2e^{3s}\end{aligned}}{e^{s}\big((w-1)e^{2s} - w\big)} = 0. \tag{4.97}$$

The denominator is not zero because we are able to solve for β. Thus, it suffices to consider the numerator which yields Formula (4.91).

We will now discuss the roots larger than 2 of this cubic polynomial. It is convenient to change variables $\theta = w - 2$, so that we are interested

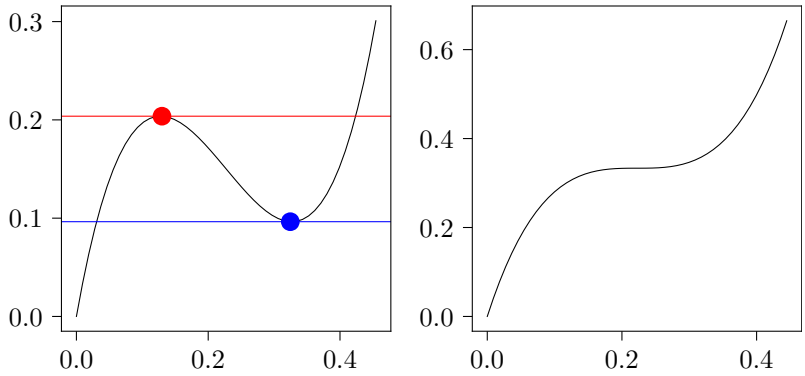

Figure 4.11: This figure shows how $v(m, \beta, t)$ behaves as a function of the y-coordinate of m for $g_t \approx 0.07012$. In the left plot ($\beta \approx 2.6685$) you see that there is a region for $v(m, \beta, t)$ such that there exist three solutions to the equation $v(m, \beta, t) = v_0$. In the right plot ($\beta \approx 2.7267$) this region is gone. For any v_0 in this region, we find three zeros of the partial derivative of the potential with respect to the y-coordinate of m corresponding to two local minimizers and a saddle point. The red and blue dots correspond to the same dots in the central plot of figure 4.10

in the positive roots of the following polynomial:

$$\begin{aligned} \theta^3(e^{3s} - e^s) &- (6se^{2s} + e^{4s} - 4e^{3s} - 3e^{2s} + 5e^s - 2)\theta^2 \\ &- (18se^{2s} + 2e^{4s} - 7e^{3s} - 9e^{2s} + 10e^s - 8)\theta \\ &- 12se^{2s} - e^{4s} + 4e^{3s} + 6e^{2s} - 8e^s + 8 \end{aligned} \tag{4.98}$$

Using Descartes' rule of signs, we know that the number of positive roots is equal to the number of sign changes among consecutive, nonzero coefficients of the polynomial or it less than it by an even number. Note that the coefficients in increasing order for $s = 0$ are given by $(9, 12, 3, 0)$. Therefore we do not find any positive roots for very low positive values of s. The first sign changes appears for the coefficient of order zero which yields equation (4.92). All of the coefficients except the highest order coefficient eventually become negative. However, with increasing s this happens with increasing order of the coefficient so that we have only one sign change between consecutive coefficients for each s larger than s_*. Thus, for all $s > s_*$ there exists only one root $w_*(s)$ larger than 2. □

4.4.8 Reentry into Gibbs: the Maxwell triangle exit (MTE) line

For β in the interval $(\frac{8}{3}, 4\log 2)$ the model displays recovery as well but due to a different mechanism. After the horns of two pentagrams have touched, the Maxwell set which consisted of three connected components now has become one connected component. It consists of three straight lines on the axes of symmetry and a triangle with curved edges. The model recovers from the non-Gibbsianness when this triangle completely leaves the unit simplex which happens on another line in the dynamical phase diagram we call *Maxwell triangle exit (MTE)*.

Proposition 45. *For any β in the interval $(\frac{8}{3}, 4\log 2)$ define the function*

$$w(\beta, y) = 1 + \frac{(\beta + 6y)e^{-3y}}{\beta - 3y}. \tag{4.99}$$

The Maxwell triangle leaves the simplex at $t = t_{\mathrm{MTE}}(\beta) = \frac{1}{3}\log\frac{w(\beta,y)+1}{w(\beta,y)-2}$ where y in $(-\frac{\beta}{6}, \frac{\beta}{3})$ is such that there exists a y' in $(-\frac{\beta}{6}, \frac{\beta}{3})$ and (y, y') is a solution of the system

$$(\beta + 6y)(\beta - 3y')e^{-3y} - (\beta + 6y')(\beta - 3y)e^{-3y'} = 0 \tag{4.100}$$

$$-2y - y' - \frac{3}{\beta}\Big((y')^2 - y^2\Big) + \log\frac{\beta}{3}\left(-2\,(\beta - 3\,y)e^{3\,y} - (\beta + 6\,y)e^{3\,y'}\right) = 0 \tag{4.101}$$

Before we come to the proof, let us remark the following: Of course, it is impractical to solve this system by hand. However, for fixed β we can show the zeros of the left-hand sides of both equations. Figure 4.12 shows them in the relevant rectangle $(-\frac{\beta}{6}, \frac{\beta}{3}) \times (-\frac{\beta}{6}, \frac{\beta}{3})$. The line as depicted in the dynamical phase diagram is obtained via a numerical solution of this system of equations.

Proof. Let $m = \varphi_\beta^{-1}(0, y, 0)$ be any point on the axis of symmetry with $m_2 = m_3$. This point is mapped to $\alpha = (1, 0, 0)$ by the catastrophe map

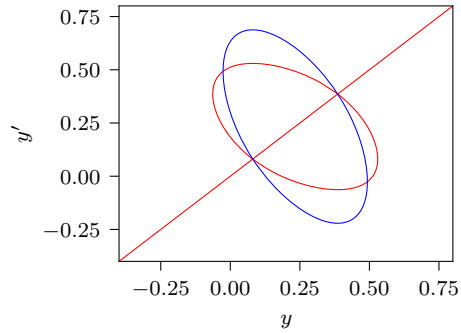

Figure 4.12: The zeros of the left-hand sides of the two equations (4.100) and (4.101) for $\beta = 2.8$. The red curve corresponds to the solutions of (4.100) and the blue curve to the solutions of (4.101). The intersection of the red curve with diagonal is of course a trivial solution and not the one we are looking for.

$\chi(\cdot, \beta, t)$ if and only if

$$\frac{6y}{\beta} + 1 - \frac{3(w-1)e^{3y}}{(w-1)e^{3y}+2} = 0 \tag{4.102}$$

which is the equation $\frac{\partial G_{\alpha,\beta,t}}{\partial y} = 0$ where $\alpha = (1,0,0)$ and we have substituted $w = e^{g_t} + 1$. Solving this equation for w we find two solutions one of which is positive. This yields (4.99).

Let $m' = \varphi_\beta^{-1}(0, y', 0)$ be any point on the same axis of symmetry. The value of $G_{\alpha,\beta,t}$ at these two points m and m' are equal if and only if $G_{\alpha,\beta,t} - G_{\alpha,\beta,t} = 0$. Plugging in $t = \frac{1}{3}\log\frac{w(\beta,y)+1}{w(\beta,y)-2}$ and $\alpha = (1,0,0)$ yields (4.101). Equation (4.100) comes from the fact that m and m' are stationary points that belong to the same time variable t, that is, $w(\beta, y) - w(\beta, y') = 0$. If we multiply this equation by $(\beta - 3y)(\beta - 3y')$ we arrive at (4.100). □

4.5 Loss of the Gibbs property without recovery

If β lies in the interval $(4\log 2, 3)$, the model displays the loss of the Gibbs property without recovery. This is due to the uniform distribution which becomes bad after a sharp transition time and stays bad forever.

This behavior is analogous to the behavior in the static model described by the Ellis-Wang theorem [14].

4.5.1 The Ellis-Wang (EW) line

The static model has a phase-coexistence of four states at inverse temperature $4\log 2$ in zero field [14]. The first layer model as discussed in this paper has a whole line of such points which we refer to as Ellis-Wang points.

Proposition 46. *Suppose $\alpha = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, that is, it represents the uniform distribution.*

(a) The HS transform $G_{\alpha,\beta,t}$ has a point of phase-coexistence with four global minimizers if and only if there exists a solution (s, β, t) to the following system of equations.

$$\frac{3y}{\beta} + \frac{1}{e^{3y+g_t}+2} - \frac{e^{3y}}{e^{3y}+e^{g_t}+1} = 0 \tag{4.103}$$

$$3y\left(1+\frac{3y}{\beta}\right) + \log\frac{(e^{g_t}+2)^3}{(e^{g_t}+1+e^{3y})^2(e^{3y+g_t}+2)} = 0 \tag{4.104}$$

(b) The solutions to the above systems can be parametrized in terms of $s = 3y$ given via

$$\beta = \frac{s\big(e^s(w_*(s)-1)+2\big)(e^s+w_*(s))}{(e^s-1)(w_*(s)e^s+w_*(s)-e^s)}, \tag{4.105}$$

$$g_t = \log(w_*(s)-1) \tag{4.106}$$

where $s > 2\log 2$ and $w_*(s)$ is the unique zero in $(2,\infty)$ of

$$w \mapsto s\left(1+\frac{(e^s-1)(we^s-e^s+w)}{(w+e^s)(we^s-e^s+2)}\right)+\log\frac{(w+1)^3}{(w+e^s)^2(we^s-e^s+2)}. \tag{4.107}$$

Proof. First, let us derive the system of equations (4.103–4.104). Since α has the full symmetry, that is, it is invariant under any permutation of S_3, it suffices to consider the equal-depth of the central minimum m_0

with one of the three outer ones denoted by m. In the following, we assume $m_2 = m_3$. The relative difference between the values is given by

$$G_{\alpha,\beta,t}(m) - G_{\alpha,\beta,t}(m_0) = y + \frac{3y^2}{\beta} - \frac{1}{3}\log\Big(e^{g_t+3y} + 2\Big) - \frac{2}{3}\log\Big(e^{g_t} + e^{3y} + 1\Big) + \log(e^{g_t} + 2). \tag{4.108}$$

By collecting the logarithmic terms and multiplying the equation by 3 we find (4.104). Equation (4.103) comes from the fact that m is a stationary point. So we calculate the relevant partial derivative

$$\frac{\partial G_{\alpha,\beta,t}}{\partial y} = \frac{6y}{\beta} + 1 - \Gamma_{1,1} - 2\Gamma_{2,1} \tag{4.109}$$

where $\Gamma_{b,a} = \Gamma_{b,a}(\beta m, t)$. The partial derivative with respect to the x-coordinate of m vanishes because of symmetry. Plugging in the expressions for $\Gamma_{1,1}$ and $\Gamma_{2,1}$ yields (4.103).

Now, let us come to the parametrization. Equation (4.105) follows by substituting $w = e^{g_t} + 1$ and $s = 3y$ in equation (4.103) and solving for β which is possible since $s \neq 0$. Plugging this into equation (4.104) and making the same substitutions we find (4.107). Note that $w_*(s)$ is increasing with s and that the solution of $w_*(s) = 2$ is $s = 2\log 2$. For lower values of s (4.107) has no zeros larger than two. □

4.5.2 The elliptic umbilics (EU) line

In the static model there is a special point called elliptic umbilic. This catastrophe at the center of the unit simplex is responsible for the fact that the central minimum changes to a maximum. In the dynamical model – due to the additional parameter g_t – we have a whole line of these points. This line we call the line of elliptic umbilics (EU).

Proposition 47. *For each $\beta \geq 3$ define the function*

$$w(\beta) = \beta - 1 + \sqrt{\beta(\beta - 3)}. \tag{4.110}$$

Fix some $\beta \geq 3$ and let $\alpha = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $t = \frac{1}{3}\log\frac{w(\beta)+1}{w(\beta)-2}$. Then:

(a) The Hessian $G''_{\alpha,\beta,t}(m)$ at $m = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ has a double zero eigenvalue.

(b) The Taylor expansion of $G_{\alpha,\beta,t}$ at $m = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ for $\beta = 3$ (and therefore $g_t = 0$) up to the third order is given by

$$x^2 y - \frac{1}{3}y^3 + \frac{1}{2}z^2 - \log 3 - \frac{1}{2}. \tag{4.111}$$

Proof. First, we check that the Hessian has a double zero eigenvalue. Let α equal $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and consider the Hessian of $G_{\alpha,\beta,t}$ at $m = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. With the same arguments as in the proof of Proposition 37, we find that the Hessian is diagonal. Furthermore, since α and m have the full symmetry, the two second order partial derivatives $\frac{\partial^2 G_{\alpha,\beta,t}}{\partial y^2}$ and $\frac{\partial^2 G_{\alpha,\beta,t}}{\partial x^2}$ are equal. Let us consider the partial derivative with respect to y.

$$\begin{aligned}
\frac{\partial^2 G_{\alpha,\beta,t}}{\partial y^2} &= \frac{6}{\beta} - 3\Big(\Gamma_{1,1} - \Gamma_{1,1}^2 + 2(\Gamma_{2,1} - \Gamma_{2,1}^2)\Big) \\
&= \frac{6}{\beta} - 3\left(\frac{e^{g_t}}{e^{g_t}+2} - \frac{e^{2g_t}}{(e^{g_t}+2)^2} + \frac{2}{e^{g_t}+2} - \frac{2}{(e^{g_t}+2)^2}\right) \\
&= \frac{6}{\beta} - 3\left(1 - \frac{(w-1)^2+2}{(w+1)^2}\right) \\
&= 6\frac{w^2 + 2(1-\beta)w + 1 + \beta}{\beta(w+1)^2}
\end{aligned} \tag{4.112}$$

where $\Gamma_{b,a} = \Gamma_{b,a}(\beta m, t)$ and we have substituted $w = e^{g_t} + 1$. Setting this equal to zero and solving for w yields (4.110) since the other root of the quadratic polynomial in the numerator is always less than two.

Now we come to (b). Plugging $\beta = 3$ and $g_t = 0$ into the HS transform and writing it in the (x, y, z)-coordinates we arrive at

$$\begin{aligned}
G_{\alpha,\beta,t}(m) = \frac{3}{2}\langle m, m\rangle - \log \sum_{a=1}^{3} e^{3m_a} = x^2 + y^2 + \frac{1}{2}z^2 + \sqrt{3}x \\
+ y - \frac{1}{2} - \log\Big(1 + e^{2\sqrt{3}x} + e^{\sqrt{3}x + 3y}\Big).
\end{aligned} \tag{4.113}$$

Using the Taylor expansion of the logarithm and the exponential function, (4.111) follows by an elementary computation. Note that (4.113) is actually the HS transform of the static Potts model. □

Using symbolic computation with the help of a computer, it is also possible to obtain a Taylor expansion for every pair (β, g_t) on the Elliptic umbilic line. Because of symmetry, the β-dependent coefficients of x^2y and y^3 differ only by a factor of $-\frac{1}{3}$. This means that for any (β, g_t) on the Elliptic umbilic line the potential $G_{\alpha,\beta,t}$ with α representing the uniform distribution has the following Taylor expansion up to order three around the simplex center.

$$\frac{A_1(\beta)}{A_2(\beta)}\left(x^2y - \frac{1}{3}y^3\right) + \frac{3}{2\beta}z^2 - \frac{1}{6}\beta - \log\left(\beta + \sqrt{\beta(\beta-3)}\right) \tag{4.114}$$

The functions $A_1(\beta), A_2(\beta)$ are given as follows:

$$\begin{aligned} A_1(\beta) = {} & 7077888\,\beta^{10} - 107937792\,\beta^9 + 700710912\,\beta^8 - 2523156480\,\beta^7 \\ & + 5502422016\,\beta^6 - 7445737728\,\beta^5 + 6152433408\,\beta^4 \\ & - 2930719968\,\beta^3 + 712130940\,\beta^2 - 67493007\,\beta + 1062882 \\ & + 27\,B(\beta)\sqrt{\beta(\beta-3)} \end{aligned} \tag{4.115}$$

$$\begin{aligned} B(\beta) = {} & 262144\,\beta^9 - 3604480\,\beta^8 + 20840448\,\beta^7 - 65802240\,\beta^6 \\ & + 123282432\,\beta^5 - 139366656\,\beta^4 + 92378880\,\beta^3 - 33102432\,\beta^2 \\ & + 5380020\,\beta - 255879 \end{aligned} \tag{4.116}$$

$$\begin{aligned} A_2(\beta) = {} & 1048576\,\beta^{12} - 16515072\,\beta^{11} + 111476736\,\beta^{10} - 421134336\,\beta^9 \\ & + 975421440\,\beta^8 - 1426553856\,\beta^7 + 1307674368\,\beta^6 \\ & - 720555264\,\beta^5 + 218245104\,\beta^4 - 30311820\,\beta^3 \\ & + 1240029\,\beta^2 + C(\beta)\sqrt{\beta(\beta-3)} \end{aligned} \tag{4.117}$$

$$C(\beta) = 1048576\,\beta^{11} - 14942208\,\beta^{10} + 90243072\,\beta^{9} - 300810240\,\beta^{8} + 603832320\,\beta^{7} - 747242496\,\beta^{6} + 560431872\,\beta^{5} - 240185088\,\beta^{4} + 51963120\,\beta^{3} - 4330260\,\beta^{2} + 59049\,\beta \tag{4.118}$$

Source code for symbolical computation As we have already mentioned, since the expressions showing up in the characterizations of the transition lines are long, we use the SageMath package for our symbolic computation. In fact the expressions are so long that evaluating those expressions using the Sage interpreter is very time-consuming. Therefore, we use the code generation facilities of the `sympy` library to generate C code so that we can do the numerical computation in C. We have included the file `potts-numerics-1.0.tar.gz` in the Electronic Supplemental Material (ESM) which contains the code and further information on how to use it. With this package you can, for example, create high-resolution plots of the functions involved in the computation of the lines BU, TPE and ACE, or generate bifurcation set slices and Maxwell set slices.

Bibliography

[1] V. I. Arnold, S. M. Gusein-Zade, and A. N. Varchenko. "Singularities of Differentiable Maps". In: (1985). DOI: 10.1007/978-1-4612-5154-5.

[2] S. Berghout, R. Fernández, and E. Verbitskiy. "On the relation between Gibbs and g-measures". In: *Ergodic Theory Dyn. Syst.* 39.12 (2019), pp. 3224–3249. DOI: 10.1017/etds.2018.13.

[3] S. A. Bethuelsen and D. Conache. "One-sided continuity properties for the schonmann projection". In: *J. Stat. Phys.* 172.4 (2018), pp. 1147–1163. DOI: 10.1007/s10955-018-2092-z.

[4] R. Bissacot, E. O. Endo, A. C. D. van Enter, and A. Le Ny. "Entropic repulsion and lack of the g-measure property for Dyson models". In: *Commun. Math. Phys.* 363.3 (2018), pp. 767–788. DOI: 10.1007/s00220-018-3233-6.

[5] A. Bovier and F. den Hollander. "Metastability". In: *Grundlehren der mathematischen Wissenschaften* (2015). DOI: 10.1007/978-3-319-24777-9.

[6] T. Bröcker. *Differentiable Germs and Catastrophes*. Ed. by L. Lander. London Mathematical Society Lecture Note Series. Cambridge University Press, 1975. DOI: 10.1017/CBO9781107325418.

[7] R. Cerf and M. Gorny. "A Curie–Weiss model of self-organized criticality". In: *The Annals of Probability* 44.1 (Jan. 2016), pp. 444–478. DOI: 10.1214/14-aop978.

[8] S. Chatterjee and Q.-M. Shao. "Nonnormal Approximation By Stein's Method of Exchangeable Pairs With Application To the Curie-Weiss Model". In: *The Annals of Applied Probability* 21.2 (2011), pp. 464–483. DOI: 10.1214/10-aap712.

[9] A. Dembo and O. Zeitouni. *Large deviations techniques and applications. 2nd ed., corrected 2nd printing.* English. 2nd ed., corrected 2nd printing. Vol. 38. Berlin: Springer, 2010, pp. xvi + 396.

[10] F. den Hollander, F. Redig, and W. van Zuijlen. "Gibbs-non-Gibbs dynamical transitions for mean-field interacting Brownian motions". In: *Stochastic Processes Appl.* 125.1 (2015), pp. 371–400. DOI: 10.1016/j.spa.2014.09.011.

[11] P. Eichelsbacher and B. Martschink. "On rates of convergence in the Curie–Weiss–Potts model with an external field". In: *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* 51.1 (Feb. 2015), pp. 252–282. DOI: 10.1214/14-aihp599.

[12] R. S. Ellis. *Entropy, large deviations, and statistical mechanics. Reprint of the 1985 edition.* English. Reprint of the 1985 edition. Berlin: Springer, 2006, pp. xiv + 364.

[13] R. S. Ellis and C. M. Newman. "Limit theorems for sums of dependent random variables occurring in statistical mechanics." In: *Z. Wahrscheinlichkeitstheor. Verw. Geb.* 44 (1978), pp. 117–139. DOI: 10.1007/BF00533049.

[14] R. S. Ellis and K. Wang. "Limit theorems for the empirical vector of the Curie-Weiss-Potts model." In: *Stochastic Processes Appl.* 35.1 (1990), pp. 59–79. DOI: 10.1016/0304-4149(90)90122-9.

[15] V. Ermolaev and C. Külske. "Low-temperature dynamics of the Curie-Weiss model: Periodic orbits, multiple histories, and loss of Gibbsianness". In: *J. Stat. Phys.* 141.5 (2010), pp. 727–756. DOI: 10.1007/s10955-010-0074-x.

[16] R. Fernández, F. Den Hollander, and J. Martínez. "Variational description of Gibbs-non-Gibbs dynamical transitions for spin-flip systems with a Kac-type interaction". In: *J. Stat. Phys.* 156.2 (2014), pp. 203–220. DOI: 10.1007/s10955-014-1004-0.

[17] R. Fernández, F. den Hollander, and J. Martínez. "Variational description of Gibbs-non-Gibbs dynamical transitions for the Curie-Weiss model". In: *Commun. Math. Phys.* 319.3 (2013), pp. 703–730. DOI: 10.1007/s00220-012-1646-1.

[18] R. Fernández, F. den Hollander, and J. Martínez. "Variational Description of Gibbs-non-Gibbs Dynamical Transitions for the Curie-Weiss Model". In: *Communications in Mathematical Physics* 319.3 (Dec. 2012), pp. 703–730. DOI: 10.1007/s00220-012-1646-1.

[19] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems. A concrete mathematical introduction*. English. Cambridge: Cambridge University Press, 2018, pp. xix + 622.

[20] J. Gaite, J. Margalef-Roig, and S. Miret-Artés. "Analysis of a Three-Component Model Phase Diagram By Catastrophe Theory". In: *Physical Review B* 57.21 (1998), pp. 13527–13534. DOI: 10.1103/physrevb.57.13527.

[21] J. Gaite, J. Margalef-Roig, and S. Miret-Artés. "Analysis of a Three-Component Model Phase Diagram By Catastrophe Theory: Potentials With Two Order Parameters". In: *Physical Review B* 59.13 (1999), pp. 8593–8601. DOI: 10.1103/physrevb.59.8593.

[22] J. A. Gaite. "Phase Transitions As Catastrophes: The Tricritical Point". In: *Physical Review A* 41.10 (1990), pp. 5320–5324. DOI: 10.1103/physreva.41.5320.

[23] H.-O. Georgii. *Gibbs measures and phase transitions. 2nd extended ed.* English. 2nd extended ed. Vol. 9. Berlin: de Gruyter, 2011, pp. xix + 545.

[24] R. Gheissari, C. M. Newman, and D. L. Stein. "Zero-Temperature Dynamics in the Dilute Curie–Weiss Model". In: *Journal of Statistical Physics* 172.4 (June 2018), pp. 1009–1028. DOI: 10.1007/s10955-018-2087-9.

[25] O. Häggström and C. Külske. "Gibbs properties of the fuzzy Potts model on trees and in mean field". English. In: *Markov Process. Relat. Fields* 10.3 (2004), pp. 477–506.

[26] F. Henning, R. C. Kraaij, and C. Külske. "Gibbs-non-Gibbs transitions in the fuzzy Potts model with a Kac-type interaction: closing the Ising gap". In: *Bernoulli* 25.3 (2019), pp. 2051–2074. DOI: 10.3150/18-BEJ1045.

[27] E. Ising. "Beitrag zur Theorie des Ferromagnetismus". German. In: *Zeitschrift für Physik* 31.1 (Feb. 1925), pp. 253–258. DOI: 10.1007/BF02980577.

[28] B. Jahnel, C. Külske, E. Rudelli, and J. Wegener. "Gibbsian and non-Gibbsian properties of the generalized mean-field fuzzy Potts-model". In: *Markov Process. Relat. Fields* 20.4 (2014), pp. 601–632.

[29] B. Jahnel and C. Külske. "Attractor properties for irreversible and reversible interacting particle systems". In: *Commun. Math. Phys.* 366.1 (2019), pp. 139–172. DOI: 10.1007/s00220-019-03352-4.

[30] B. Jahnel and C. Külske. "The Widom – Rowlinson model under spin flip: immediate loss and sharp recovery of quasilocality". In: *Ann. Appl. Probab.* 27.6 (2017), pp. 3845–3892. DOI: 10.1214/17-AAP1298.

[31] S. Kissel and C. Külske. "Dynamical Gibbs-non-Gibbs transitions in Curie-Weiss Widom-Rowlinson models". In: *Markov Process. Relat. Fields* 25.3 (2019), pp. 379–413.

[32] S. Kissel and C. Külske. "Dynamical Gibbs-non-Gibbs transitions in lattice Widom-Rowlinson models with hard-core and soft-core interactions". In: *J. Stat. Phys.* 178.3 (2020), pp. 725–762. DOI: 10.1007/s10955-019-02478-y.

[33] A. Klenke. *Wahrscheinlichkeitstheorie*. 2nd ed. Heidelberg: Springer Verlag, 2008. DOI: 10.1007/978-3-540-77571-3.

[34] R. C. Kraaij, F. Redig, and W. B. van Zuijlen. *A Hamilton-Jacobi point of view on mean-field Gibbs-non-Gibbs transitions*. Nov. 9, 2017. arXiv: 1711.03489v1.

[35] C. Külske and A. Le Ny. "Spin-flip dynamics of the Curie-Weiss model: loss of Gibbsianness with possibly broken symmetry." In: *Commun. Math. Phys.* 271.2 (2007), pp. 431–454. DOI: 10.1007/s00220-007-0201-y.

[36] C. Külske, A. Le Ny, and F. Redig. "Relative entropy and variational properties of generalized Gibbsian measures". In: *Ann. Probab.* 32.2 (2004), pp. 1691–1726. DOI: 10.1214/009117904000000342.

[37] C. Külske and D. Meißner. *Dynamical Gibbs–non-Gibbs transitions in the Curie-Weiss Potts model in the regime $\beta < 3$.* Oct. 31, 2020. arXiv: 2011.00350.

[38] C. Külske and D. Meißner. "Stable and Metastable Phases for the Curie–Weiss–Potts Model in Vector-Valued Fields via Singularity Theory". In: *Journal of Statistical Physics* 181.3 (Nov. 2020), pp. 968–989. DOI: 10.1007/s10955-020-02615-y.

[39] C. Külske and F. Redig. "Loss without recovery of Gibbsianness during diffusion of continuous spins". In: *Probab. Theory Relat. Fields* 135.3 (2006), pp. 428–456. DOI: 10.1007/s00440-005-0469-9.

[40] P. G. Lammers and M. Tassy. "Variational principle for weakly dependent random fields". In: *J. Stat. Phys.* 179.4 (2020), pp. 846–870. DOI: 10.1007/s10955-020-02538-8.

[41] C. Landim and I. Seo. "Metastability of Non-Reversible, Mean-Field Potts Model With Three Spins". In: *Journal of Statistical Physics* 165.4 (2016), pp. 693–726. DOI: 10.1007/s10955-016-1638-1.

[42] P. S. Laplace. "Memoir on the Probability of the Causes of Events". In: *Statistical Science* 1.3 (1986), pp. 364–378.

[43] J. M. Lee. "Introduction to Smooth Manifolds". In: *Graduate Texts in Mathematics* (2012). DOI: 10.1007/978-1-4419-9982-5.

[44] W. Lenz. "Beiträge zum Verständnis der magnetischen Eigenschaften in festen Körpern". German. In: *Physikalische Zeitschrift* 21 (1920), pp. 613–615.

[45] Y.-C. Lu. "Introduction to Singularity Theory with Historical Remarks". In: *Singularity Theory and an Introduction to Catastrophe Theory*. New York, NY: Springer New York, 1976, pp. 1–23. DOI: 10.1007/978-1-4612-9909-7_1.

[46] E. Olivieri and M. E. Vares. "Large Deviations and Metastability". In: (Feb. 2005). DOI: 10.1017/cbo9780511543272.

[47] L. Onsager. "Crystal statistics. I: A two-dimensional model with an order-disorder transition". English. In: *Phys. Rev., II. Ser.* 65 (1944), pp. 117–149. DOI: 10.1103/PhysRev.65.117.

[48] R. E. Peierls. “On Ising’s model of ferromagnetism”. English. In: *Proc. Camb. Philos. Soc.* 32 (1936), pp. 477–481. DOI: 10.1017/S0305004100019174.

[49] T. Poston and I. Stewart. *Catastrophe Theory and its Applications*. London etc.: Pitman Publishing Ltd., 1978.

[50] F. Redig and F. Wang. “Hamiltonian and Lagrangian for the trajectory of the empirical distribution and the empirical measure of Markov processes”. In: *J. Stat. Phys.* 157.1 (2014), pp. 182–204. DOI: 10.1007/s10955-014-1063-2.

[51] M. Shamis and O. Zeitouni. “The Curie-Weiss Model With Complex Temperature: Phase Transitions”. In: *Journal of Statistical Physics* 172.2 (2017), pp. 569–591. DOI: 10.1007/s10955-017-1812-0.

[52] A. C. D. van Enter, R. Fernández, F. den Hollander, and F. Redig. “A large-deviation view on dynamical Gibbs-non-Gibbs transitions”. In: *Mosc. Math. J.* 10.4 (2010), pp. 687–711.

[53] A. C. D. van Enter, R. Fernández, F. den Hollander, and F. Redig. “Possible loss and recovery of Gibbsianness during the stochastic evolution of Gibbs measures”. In: *Commun. Math. Phys.* 226.1 (2002), pp. 101–130. DOI: 10.1007/s002200200605.

[54] A. van Enter and A. Le Ny. “Decimation of the Dyson-Ising ferromagnet”. In: *Stochastic Processes Appl.* 127.11 (2017), pp. 3776–3791. DOI: 10.1016/j.spa.2017.03.007.

[55] A. C. D. Van Enter, V. N. Ermolaev, G. Iacobelli, and C. Külske. “Gibbs-non-Gibbs properties for evolving Ising models on trees”. In: *Ann. Inst. Henri Poincaré, Probab. Stat.* 48.3 (2012), pp. 774–791. DOI: 10.1214/11-AIHP421.

[56] A. C. D. van Enter, R. Fernández, and A. D. Sokal. “Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory”. In: *J. Stat. Phys.* 72.5-6 (1993), pp. 879–1167. DOI: 10.1007/BF01048183.

[57] A. C. D. van Enter, C. Külske, A. A. Opoku, and W. M. Ruszel. "Gibbs-non-Gibbs properties for n-vector lattice and mean-field models". In: *Braz. J. Probab. Stat.* 24.2 (2010), pp. 226–255. DOI: 10.1214/09-BJPS029.

[58] K. Wang. "Solutions of the Variational Problem in the Curie-Weiss-Potts Model". In: *Stochastic Processes and their Applications* 50.2 (1994), pp. 245–252. DOI: 10.1016/0304-4149(94)90121-x.

[59] F. Y. Wu. "The Potts Model". In: *Reviews of Modern Physics* 54.1 (1982), pp. 235–268. DOI: 10.1103/revmodphys.54.235.

Acknowledgements

First and foremost, my sincere thanks goes to my advisor Christof Külske. His intuition and expertise in the field of mathematical physics, which he has shared with me in numerous discussions, has greatly helped me with my research. I also want to thank the Deutsche Forschungsgemeinschaft (DFG) and the Fakultät für Mathematik at Ruhr-Universität Bochum for their financial support. For fruitful discussions I want to thank Richard Kraaij and Sascha Kissel.

Finally, I want to thank my family and friends for their support, especially Anika for always backing me and for her love. Thank you.

This thesis has been supported by the Research training group "High-dimensional phenomena in probability – fluctuations and discontinuity" (RTG 2131).