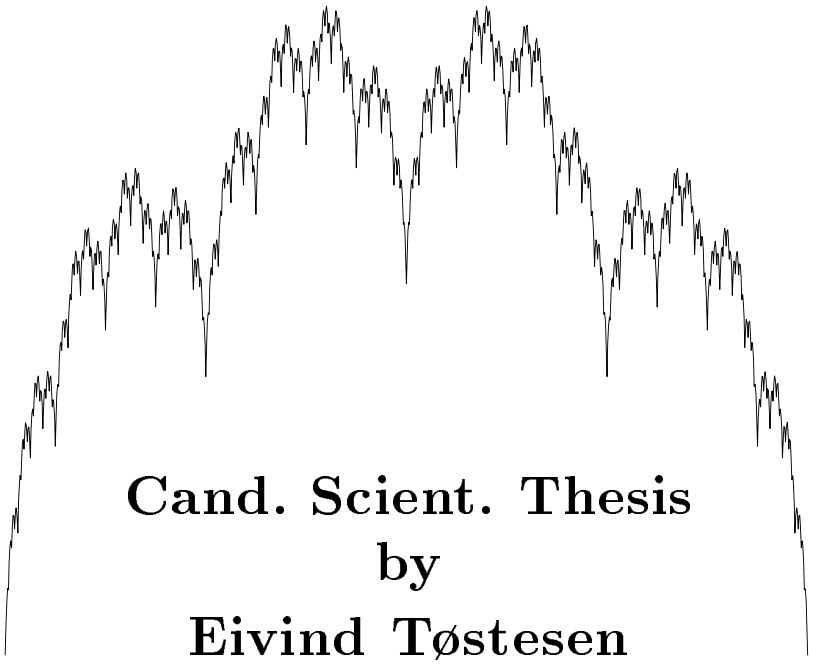
Dynamics of Hierarchically Clustered Cooperating Agents



Cand. Scient. Thesis
by
Eivind Tøstesen

University of Copenhagen
CONNECT, **The Niels Bohr Institute**
Blegdamsvej 17, 2100 Copenhagen Ø
Denmark

September 1995

Foreword

This master's thesis (or 'speciale') is the final requirement in my study for a cand. scient. degree in physics. The work was done at CONNECT (The Computational Neural Network Center) at The Niels Bohr Institute.

I would like to thank my supervisors John Hertz (Nordita), Bernardo Huberman (Xerox PARC) and Benny Lautrup (NBI) for guidance, inspiration and long discussions. John Hertz has been my main supervisor while Bernardo Huberman introduced me to the subject and supervised me during his visit fall 1993.

Thanks also to the people at CONNECT for their help. Finally a thank to the Lørup Foundation for supporting my visit in Palo Alto.

E.T.

Copenhagen, Denmark
September 1995

Contents

Preface

This thesis in physics is directed to both physics and social sciences: A model in social sciences is investigated with methods from physics. The first part, Chapters 1-3, addresses mostly questions from a social sciences point of view (dilemmas and behaviour) while the second part, chapters 4-7, is mostly inclined toward physics (dynamics and timescales).

The outline of the thesis is this: Chapter 1 is an introduction to social dilemmas and some game-theoretical background knowledge. Chapter 2 and 3 go through the theory and work of Bernardo Huberman, Natalie Glance *et al.* as presented in their articles [Outbreak] to [Multiagent]. However, it will be presented here in a different way than in the articles (for example including a mapping to Ising models), both because this is written as a thesis rather than a published article and because it is aimed at a physics audience, while the articles have a broader audience. Chapter 4 reviews background knowledge about Dyson's hierarchical model. In Chapters 5, 6 and 7 the nucleation dynamics of this model is studied using analytical and numerical techniques. Chapter 8 ties up the physics side of the models with the social science side.

Chapter 1

Social dilemmas

1.1 Introduction

A fundamental issue in social sciences (sociology, economy, political science etc.) and biology is social behaviour or cooperation on a voluntary basis. Situations abound in the world in which there is no central control to secure cooperation. It becomes an interesting problem when each individual has an incentive to be selfish and shirk from the social effort.

This is the subject of a class of problems known as *social dilemmas*. In a social dilemma each individual can choose freely whether to act for the good of the others in the group (being social, cooperative) or for the own good (being selfish, shirking). Social dilemmas are described in abstract terms via the notion of a *common good* which is produced collectively by the cooperating individuals. Cooperators contribute much to the 'size' of the common good and shirkers contribute little, or not at all. The common good is shared equally by everyone regardless of their own contribution to it. It is divisible, which means that the more people there are to share it, the less there will be for an individual. Contributing involves a cost to an individual which is greater than the benefit from the increase in the common good. An individual has therefore at any time an incentive not to contribute very much (to shirk or defect), because of the immediate gain to him. But at the same time it is more advantegeous to everyone if all cooperate than if all shirk. So conflicting choices face the individuals involved.

This abstract description applies to the following examples:

- Workers or employees of a firm/organization working in a team producing a collective output, with some kind of profit sharing for example a bonus that mirrors the team output. They all share equally because their individual efforts (how hard they work) are not monitored/supervised. So each team member is tempted to shirk and cruise at minimal effort level while getting paid for the other people's work [Training, Collaboration].

- Giving to charity or supporting grass-root movements and volunteer organizations like environmental groups, unions and political organizations involves a social dilemma. The organizations work for a good cause when they fight a common problem. But they depend on people to come and join the cooperation or give support.

- People sharing a common kitchen (for example at CONNECT[1]). Each person can choose to be asocial and not wash his/her cup after use and leave it in the sink for others to do it. But the more people shirk like this the more unpleasant it becomes to use the kitchen and the smaller is the chance of finding a clean cup. The common good in this case is a clean kitchen.

  A clean environment in general is a common good. For example, limiting air pollution constitutes a social dilemma. Pollution spreads to become a problem to everybody, but no one (only a few) wants to be a sucker who renounces the comforts of electricity, cars etc. But it pays, for example, to recycle if everybody does it.

- Sharing a computer processor: Shall I give way to higher priority programs or let all my own programs have highest priority?

- Asking questions at a lecture: A person in the audience who checks calculations on the blackboard and thinks about how to make the subject more clear has a cost. But the effort will benefit everybody in the audience.

- *The Unscrupulous Diners Dilemma* [SciAm]: A group of people are dining together in a restaurant. They are not each going to pay their

[1]Computational Neural Network Center

own meal but decide instead to divide the check evenly among them. They can order a cheap meal or an expensive meal but they don't know what the others will order. If they were paying for themselves they would order a cheap meal, considering price and quality, but because they divide the bill they now have a chance of having an expensive meal at the expense of the others. This works unless all the others do the same. Being nice to the others and ordering the modest, cheap meal is a cooperative act while ordering expensive is selfish.

- *The Tourists' Dilemma:* Imaging a group of tourists taking pictures. They all want the highly valued motives so they tend to take identical photos. It would benefit the group if they cooperated and tried to avoid this overlap. Then they could share a much more diverse pool of pictures when they return home. But this produces a social dilemma because a tourist must take pictures of lower quality if he/she tries to avoid redundancy.

  Avoiding overlap to benefit the group instead of one self is a social dilemma in many situations, not only this funny tourist story. One example is press agencies. Another is a soccer game: There are different roles that the players can play. If these roles were not distributed in advance, then maybe all the players would want to play the glorious role of the one who scores, and they all rush to the opposite goal area. But then their team would not do very well.

1.2 Game theory

The natural way to analyse these situations is to state the costs and benefits quantitatively. This is the approach of *game theory*, which is the mathematical framework we need. This approach traditionally involves a game called the Prisoner's Dilemma. I will briefly introduce some of the concepts of game theory, but I will refer to the textbooks for mathematical definitions.

In game theory a social/economic situation is described as a *game* with a number of *players*. The players must make some choices on what moves or *actions* they want to make during the game. It is specified through a *payoff* or *utility function*, giving their payoff/utility as a function of all players' actions. The payoffs are not only amounts of money but are supposed to

encompass any kind of utility. The players know the game and its payoffs and they are acting more or less intelligently or *rationally* to maximize their own payoffs. The players are assumed to be basically selfish in the sense that their only concern is their own payoffs. (A concern for other people, love, moral etc., should be included in their own payoffs.) The program of game theory is to make predictions about players' behaviour in situations that are described as games, assuming a certain rationality. This is also the program of Huberman et al. although they do not follow the traditional game theoretical approaches.

There are two kinds of game theory: cooperative and non-cooperative. We consider the non-cooperative game theory in which players can *not* commit themselves to certain future actions. It is not possible for example to make promises or agreements 'outside' the game for example before it starts. So the possibility of cheating in a social dilemma can not be circumvented this way.

The goal of analysing a game is to find *strategies*. A strategy of a player is a complete description of his behaviour (how he chooses his actions) in any situation during the game. Once a set of strategies of the players, a *strategy profile*, is found the evolution of the game is specified. And given a strategy profile the expected total payoff from the game for each player can be found.

(Non-cooperative) game theory offers a number of solution concepts. The most basic is a *Nash equilibrium*. A Nash equilibrium is a strategy profile in which every player's strategy is a best response to the other players' strategies. No player can, in other words, find an alternative strategy which will give him a higher expected utility when played against the other players keeping their Nash strategies. No player has therefore reason to change strategy and the strategies are in equilibrium. Other solution concepts are also important but they are often just refinements of the Nash equilibrium. The central idea in all of them is the equilibrium idea where each player has optimized his strategy against the other players' strategies. This is not an optimisation task against a fixed background: You do not know the other strategies that your own best answer depends on.

Game theorists calculate strategies. But the participants in games, as in real life, seldomly carry out a mathematical analysis before the game to find a strategy. Game theorists try to justify their solutions in other ways. Equilibria can come about as a result of learning or evolution, as well as intelligent thinking. A player therefore does not have to be intelligent to behave

rationally. Still, game theory has a problem with the high level of rationality assumed in the standard solution concepts. The real world is often not so perfect. Therefore game theorists are working on other 'solution' concepts that incorporate imperfectness, such as bounded rationality or irrationality, bounded complexity or capacity of players and their strategies, incomplete information and noise. Nash, Selten and Harsanyi, who have made some of the most important contributions to game theory, received the 1994 Nobel prize in economics for these developments.

1.3 Prisoner's dilemma

We use the game theoretic setting as our starting point: A social dilemma is described by the game called *the N-person Prisoner's Dilemma*: There are N players and each player has two possible actions: *Cooperate* and *defect* abbreviated C and D. There is a constant cost c for each player who cooperates and each player who cooperates adds a benefit b to the common good which is shared equally by all. The utility function u_i for player i in this game is a function of the chosen actions of all the players, but it only depends on the number n_c of players who choose C, and it is

$$u_i = \frac{n_c}{N} b - k_i c \tag{1.1}$$

where $k_i = 1$ if i cooperates and $k_i = 0$ if i defects. We require that no matter how many other players are cooperating defect will always have a higher utility so that

$$\frac{n_c + 1}{N} b - c < \frac{n_c}{N} b; \tag{1.2}$$

in other words, the marginal benefit is less than the cost of cooperation. On the other hand we require that overall cooperation is better than overall defection

$$\frac{N}{N} b - c < 0. \tag{1.3}$$

All in all,

$$b > c > \frac{b}{N} \tag{1.4}$$

The 2-person version of this game is called *the Prisoner's Dilemma*. The reason for this name is a story that goes along with the description of the

game about two prisoners in separate cells[2]. But it only diverts the attention from the generality of the situation the game describes. The Prisoner's Dilemma is the simplest game that catches the essence of the social dilemmas. One can compare its role in game theory and social science with the role a mass on a spring (harmonic oscillator) plays in mechanics and physics: a simple textbook example which is applicable 'everywhere'. The literature that involves the prisoner's dilemma is enormous and work with this game as a basis is still going on. Therefore people assign a great significance to it. I will in the following describe some of the work that has been done, just to give some background knowledge, but I am not able to give a perfect overview of the history and the state of the field.

The reason why the N-person prisoner's dilemma is designated 'dilemma' is that if it is a *one-shot* game (the game is played only once, the players choose an action each, the game ends and they receive their payoffs) then the outcome is that all players defect since defection will always give the highest payoff. Mutual defection is also the unique Nash equilibrium in the one-shot game. The dilemma or problem for the players is that if all had cooperated instead they would all be better off. The solution of this problem is to play the game more than once. Mutual cooperation then becomes possible and sustainable because a defection in a group of cooperators no longer is so attractive. A player's transition to defection will give him an immediate rise in payoff, but it may cause the others to defect too in the future with a following drop in payoff.

1.3.1 The repeated game

The *Repeated* N-person Prisoner's Dilemma is a game that is played in discrete time. At each time step the players play the N-person prisoner's dilemma game, which in this case is called the *stage game* of the repeated game. The players know each others actions in the past with a memory of infinite or finite range. A strategy in the repeated game is a specification of the player's action at any time step given as a function of the past history. At

[2]They are suspected to have comitted a crime together. They are each given the choice of confessing or not. If both confess they will be punished hard. If both refuse to confess they will get a mild sentence. If only one of them confesses then he will be released with a reward while the other will be punished very hard. Confess=Defect and Not Confess=Cooperate. (From [Lindgren]).

time t when the players must decide which action to take in the stage game, they should not shortsightedly optimize the outcome of that stage game. That would just give mutual defection again. Instead they must also take into account payoffs in the remainder of the repeated game. They should want some sum of both the present *and* future payoffs to be as big as possible. Most commonly future payoffs are *discounted* in that sum: Let $u_i(t')$ be the payoff to player i at future time step t'. (This depend on the future actions of all the players and therefore on their strategies. If the strategies are not deterministic we must talk about expected future actions and payoffs). Player i must then at time t optimize[3] the sum

$$U_i(t) = \sum_{t'=t}^{\infty} \delta^{t'-t} u_i(t') \tag{1.5}$$

This is the expected utility accumulated in the remaining game. δ is the *discount parameter* and it is positive and less than one. Payoffs that are discounted are given lesser value the farther away they are in the future. This may reflect rates of interest or that the future payoffs are unreliable, for example if at any time step there is a probability $1-\delta$ that the repeated game stops. Alternatives to the discounted sum exist, for example

$$U_i(t) = \lim_{T\to\infty} \frac{1}{T} \sum_{t'=t}^{t+T} u_i(t') \tag{1.6}$$

but are rarely used. With a stop probability the average length of the game becomes

$$H = \frac{\delta}{1-\delta} \tag{1.7}$$

which is called the *horizon*. Each player has a horizon which is his *belief* about how long he expects the game to last. For convenience each player has the same horizon. It is not the number of times the game is played. If the game were repeated a finite number H of times then backwards induction gives defection at all time steps: At time step H the players have no future and are left with a one-shot game. This means defection at the last time step. At time $H-1$ the players do not have to care about the future because they already know that there will be defection. This makes also the stage game at

[3]Remember that what is optimal depends on what the other players are doing

time step $H-1$ effectively a one-shot game with defection as the outcome, and so on back to the first time step. So we must have a probability δ that the game continues or that it goes on forever. Then a *folk theorem* states that cooperation is possible. It says that there exists a Nash equilibrium (actually a 'subgame perfect equilibrium', which is a stronger statement) in the repeated game if δ is big enough, in which the outcome is that everybody cooperates at each time step. The strategies of the players are simply to cooperate if everybody cooperated in the past and to defect otherwise. If someone uses another strategy and tries to defect he will be punished: He loses the future common good because all turn into defectors. That δ has to be big enough means that future returns must be given enough weight if the retaliation is to be felt. Unfortunately the folk theorem does not help us in predicting strategies. In fact equilibria also exist where the outcome is that a certain number of players always cooperate and the rest always defect. And many other equilibria exist. This is one of the problems with the standard solution concepts of game theory.

1.3.2 Evolution

A famous and frequently cited work is the following. Axelrod ran a tournament for computer programs. A number of scientists were invited to submit programs to play the two person prisoner's dilemma repeatedly against the same opponent. All the programs played against each other and their total payoffs were evaluated. The winning program (or strategy) was *TIT FOR TAT* submitted by Anatol Rapoport. This strategy simply plays C in the first move and then does whatever the opponent did in the previous move. If its opponent defected it also defects and if its opponent cooperated it also cooperate. Axelrod then ran an evolutionary system where the frequency of a program in the population was in proportion to its total payoff in the previous round. Here TIT FOR TAT won again: It took over the whole population. Evolutionary systems like that have become an important part of game theory: evolutionary game theory, [MaynardSmith].

These results help explain how cooperation and altruism can arise in biological evolution [Axelrod, Altruism]. Altruism or helping other individuals genes to reproduce may seem problematic, since traits of an individual that help its own genes are selected. Obviously it only pays to be altruistic if the other part is helping in return. In many examples this is automatically

the case, they both can only gain from the cooperation (mutualism). But in other cases the favour in return is not guaranteed, because the help is returned at a later time (reciprocity). A mechanism like TIT FOR TAT can explain the evolution of this.

A note of guidance to the literature on altruism: Biological altruism is very common where the help is *not* returned, for example parents helping their offspring. This is explained by *kin selection*: An individual can help its own genes by helping close relatives, because the genes have a high probability of being present in them too.

Many people have continued to simulate evolution of strategies in the iterated two person prisoner's dilemma, for example to study evolutionary dynamics in general [Lindgren] or cooperation among states internationally [Lomborg].

1.4 Finer details of cooperation and utilities

Social dilemmas deal only with questions like: Will the participants in a cooperation shirk? Will collective action spontaneously emerge? Shirking is a question of trust, punishment and retaliation. In case others shirk one must be able to cancel the cooperation quickly. Emergence of cooperation is a question of the risk in trying to initiate cooperation. Others may not follow your example. But we neglect the question of what cooperation consists in. What are the details of the collaborative work? Issues like communication, signals, planning, coordination, negotiation, and deals are aspects of cooperation out of our focus of attention. If the action C in the N-person prisoner's dilemma covers a process involving interactions in finer detail between cooperators before the result is produced then social dilemma theory view things on a coarse super level. If the game is played more than once then the time scale on which players choose between C and D should be larger than the time scale on which the detailed interactions of the process goes on, so that the result and the payoff can be produced before they reconsider their choices. Phenomena that happen on one timescale (or something-else-scale) will be seen as a 'uniform background' on other timescales.

The utility of the common good of course depends on these finer details. This leads us to the question of other utility functions than eq. 1.1. This function describes the examples listed in the beginning of the chapter fairly

well. I will now list some examples of social dilemmas with other utility functions that could be more appropriate than eq. 1.1.

- Collaborative problem solving with exchange of hints between agents [Collaboration, Micro]. Here the common good (=team performance = the speed of finding a solution) may not increase linearly with the number of cooperators as it does in eq. 1.1 but rather something between a linear and a quadratic increase, so the utility function is

$$u_i = b\frac{n_c^{1+\epsilon}}{N} - k_i c \tag{1.8}$$

- Computational societies and coevolutionary systems [Multiagent]. It is believed that social dilemmas commonly arise in situations where individual optimal choices of the agents takes the system to a situation where everybody is worse off. Such dilemmas may arise in subtle ways and may not be apparent to the system designer. Such social dilemmas may counterintuitively arise as a result of adding resources to the agents.

- Collective action against a suppressive force, for example desertion from the military, revolution or resistance in an occupied country, tax evasion, illegal public demonstrations [SciAm]. There is a police force with a certain capacity so that the risk of arrest, which is the cost of joining (cooperate), is aN/n_c. In tax evasion or desertion the personal benefit from joining the deserters would be t (='tax') and the benefit to all in having a state or military could be linear in the number of non-deserters: $b(1 - n_c/N)$

$$u_i = b(1 - \frac{n_c}{N}) + k_i(t - a\frac{N}{n_c}) \tag{1.9}$$

  In resistance work or in a demonstration the benefit to all increases with the number who have joined but now there is an additional personal cost $k_i c$

$$u_i = b\frac{n_c}{N} + k_i(-c - a\frac{N}{n_c}) \tag{1.10}$$

- Supporting science or public radio: Here we do not need to divide the cake. Information can be accessed by arbitrarily many. So the $1/N$ in eq. 1.1 is not needed:
$$u_i = n_c b - k_i c \tag{1.11}$$

- Spread of communication technology. The cost of having a telephone is $k_i c$ while the benefit in having it is proportional to the number n_c to whom you can talk.
$$u_i = k_i(n_c b - c) \tag{1.12}$$

- Buying green ecological products.

- Cooperation among states, as in the European Union.

- Slowing down military arms races or the population explosion.

If we shall keep the characteristics of social dilemmas described in the beginning (two actions, an incentive to shirk at any time, an increasing benefit with increasing number of contributors and an utility that only depend on n_c) then the utility functions of the N players must have the general form

$$u_i = \begin{cases} f_{N,i}(n_c) & if \quad i \ \ defects \\ g_{N,i}(n_c) & if \quad i \ \ cooperates \end{cases} \tag{1.13}$$

so that u_i depends on n_c, the number N and i's own choice. The functions $f_{N,i}$ and $g_{N,i}$ must be increasing and for all n_c:

$$f_{N,i}(n_c + 1) < g_{N,i}(n_c) \tag{1.14}$$

The game is invariant with respect to changes of the utility scale and a fixed additive utility i.e. if we perform the transformation

$$\begin{aligned} f_{N,i} &\to a f_{N,i} + b \\ g_{N,i} &\to a g_{N,i} + b \end{aligned} \tag{1.15}$$

for all N. The N-person prisoner's dilemma has $f_{N,i}(n_c) = bn_c/N - c$ and $g_{N,i}(n_c) = bn_c/N$. We will only use this utility function in the following and not make use of generalisations or alternatives.

Chapter 2

The flat model

In this chapter I describe the theory of social dilemmas as derived in [Outbreak] or [Diversity] and in [Beliefs] for a homogeneous, structureless (flat) group.

2.1 The rules of the game

A group of N *agents* is engaged in a social dilemma. We use the word 'agents' instead of 'group members', 'players', 'people', 'participants', 'individuals', 'computational agents', 'animals', 'organisations', 'employees' or whatever they are. The agents must make decisions on what actions to take in order to ensure a best utility for themselves, and these decisions must be made based on their information or knowledge. They know of course the rules of the game, and they should have information about the past history of the game. We want the theory to incorporate that this information may be *imperfect*. It may be uncertain, incomplete and inconsistent. Furthermore information about other parts of the system of agents is often *delayed*. We also want the theory to incorporate *noise*: Information may be noisy and agents may make mistakes. (This may reflect bounded rationality or unforeseen obstacles). Also systems with no central controller are often not synchronized. In repeated games in discrete time the players choose actions in a synchronized fashion. But we want an asynchronous process in *continuous* time (see [cellular] for a discussion).

With continuous time the utility comes in a continuous flow to each agent.

Our utility function

$$u_i(t) = \frac{b}{N}\hat{n}_c(t) - c\hat{k}_i(t) \tag{2.1}$$

now indicates the *rate* of this flow (utility per time unit) to agent i. Again $\hat{n}_c(t)$ (or $\hat{n}_c$) is the number of cooperating agents at time t and $\hat{k}_i(t)$ or $\hat{k}_i$ indicates if agent i cooperates at time t ($\hat{k}_i = 1$ cooperate $\hat{k}_i = 0$ defect). The state of the system of agents at time t is $\hat{k} = (\hat{k}_1, \ldots, \hat{k}_N)$. The number of cooperators is $\hat{n}_c = \sum_{i=1}^{N} \hat{k}_i$.

The asynchronicity is achieved by letting the agents choose at different times determined by a Poisson process. Each agent has an updating mechanism that becomes alive once in a while. This happens at rate α, and the waiting time t between successive updates is distributed like $P(t > l) = e^{-\alpha l}$. When an agent updates he can choose a new action or stay put. The rate at which the system updates is $N\alpha$ (apply the calculation eq. A.14). Between updates their actions are constant. Already now we see that one can not predict the future state of the system. It has stochastic dynamics and one can at most predict a distribution $P_{\hat{k}}(t)$ of states, and averages.

As we had in the repeated/iterated game we must also here have that agents consider their future prospects. The expected utility for the remaining game with discounting (eq.1.5) now becomes

$$U_i(t) = \int_t^\infty \langle u_i(t')\rangle e^{-\frac{(t'-t)}{H}} dt' \tag{2.2}$$

where $\langle u_i(t')\rangle$ is the rate of utility to agent i as he expects it to be at time t' in the future, and the horizon H again is the mean lifetime of the game if we interpret $e^{-\frac{(t'-t)}{H}}$ as a survival probability of the game. We shall see later how we model the agents' expectations about the future, but in general agent i at time t must have expectations in the form of probabilities over the system states as a function of time: $P_{\hat{k}}(t')$. Then his rate of utility as he expects it to be at time t' in the future is

$$\begin{aligned} \langle u_i(t')\rangle &= \sum_{\hat{k}} P_{\hat{k}}(t')(\frac{b}{N}\hat{n}_c - c\hat{k}_i) \\ &= \frac{b}{N}\langle \hat{n}_c\rangle - c\langle \hat{k}_i\rangle \end{aligned} \tag{2.3}$$

where $\langle \hat{n}_c\rangle$ and $\langle \hat{k}_i\rangle$ are the expectation values of the future number of cooperators and his own indicator.

The task for agent i if he updates at time t is to choose his action (C or D) in a way he thinks will give him the best expected payoff $U_i(t)$. This decision must be made on the basis of his knowledge of the past history. We learned that this makes cooperation possible. It is part of the game that their collectively produced common good is shared equally by all. They do not know who contributed and who did not, so they only know about $\hat{n}_c$ (this may for example be deduced from the received utility). They do not know the state $\hat{k}$.

We have delayed information in the model like this: An agent has information at time t about the number $\hat{n}_c(t-\tau)$ of cooperators at the earlier time $t-\tau$. The delay is τ. We assume the same delay for everyone. Agents have no other information. If the delay is small, $\tau \ll 1/N\alpha$, then the probability that the system has changed its state in the delay time interval is small and we might as well use $\hat{n}_c(t)$ instead of $\hat{n}_c(t-\tau)$ i.e. have zero delay.

We will model noise and imperfect information in two alternative ways: Model A and model B. Both models use the idea that noise may turn C into D with probability $1-p$ and it may turn D into C with probability $1-q$. Usually we just have $q=p$. Let $\hat{f}_c(t)$ denote the *fraction* of agents cooperating at time t: $\hat{f}_c(t) = \hat{n}_c(t)/N$.

Model A: Agent i, updating at time t, knows as input the value of $\hat{n}_c(t-\tau)$. So the perceived number and the true number of cooperators are identical in model A. He then makes a decision and forms an intention k_i to either cooperate ($k_i = 1$) or defect ($k_i = 0$). If he intends to cooperate then his actual action will be C with probability p and D with probability $1-p$. And if he intends to defect his action will be D with probability q and C with probability $1-q$. His action may be different from his intention because of noise. This agent model is sketched in fig. 2.1. The number of agents intending to cooperate at time t is denoted by $n_c(t)$ and is equal to $\sum_{i=1}^{N} k_i(t)$ and the fraction of agents intending to cooperate is $f_c = n_c/N$. What are the probabilities of different $\hat{n}_c$-values if n_c is given? $\hat{n}_c$ is the sum of two terms: The number S_p of agents who intended to cooperate and did so succesfully and the number S_q of agents who intended to defect but failed in their attempt. S_p and S_q are binomially distributed:

$$\begin{aligned} S_p &\in B(n_c, p) \\ S_q &\in B(N-n_c, 1-q) \end{aligned} \tag{2.4}$$

From this one can find the distribution of $\hat{n}_c = S_p + S_q$.

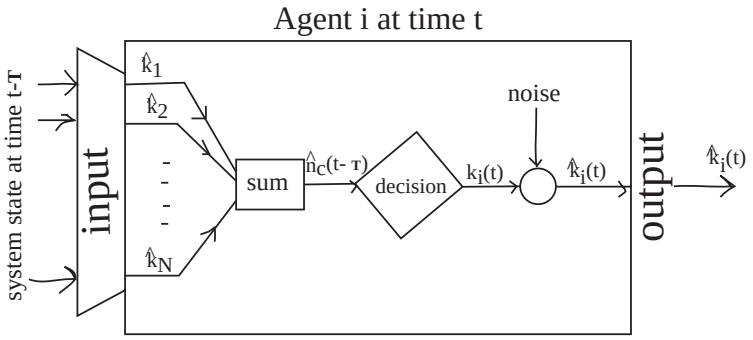


Figure 2.1: Model A

Model B: When agent i at time t is updating he receives information about $\hat{n}_c(t-\tau)$. Then this information is exposed to noise. He will *perceive* a number of cooperators $n_c^i(t-\tau)$ which is different from the actual number $\hat{n}_c(t-\tau)$. An agent that cooperates will be perceived as cooperating with probability p and as defecting with probability $1-p$ and a defecting agent will be perceived as defecting with probability q and as cooperating with probability $1-q$. The perceived number $n_c^i(t-\tau)$ will therefore be a stochastic variable $n_c^i = S_p + S_q$ where again

$$\begin{aligned} S_p &\in B(\hat{n}_c, p) \\ S_q &\in B(N-\hat{n}_c, 1-q) \end{aligned} \tag{2.5}$$

This agent model is sketched in fig. 2.2 The agent bases the decision on the

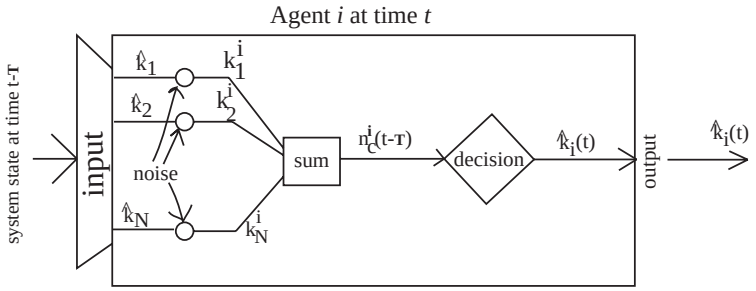


Figure 2.2: Model B

perceived number. In this model the intentions and the actual actions are

identical. We denote the action of agent j as perceived by agent i by k_j^i and the perceived fraction is

$$f_c^i = \frac{n_c^i}{N} = \frac{1}{N}\sum_{i=1}^{N} k_j^i \tag{2.6}$$

Both models have the round noise-box sitting only one place if we follow the 'wire' from agent i to agent j. But in model A it sits in agent i and acts when agent i updates while in model B it sits in agent j and acts when agent j updates. In model B there is much more noise: A decision coming out of an agent's decision-box is exposed to noise about N times as often in the system using model B than using model A.

In [Outbreak], [Beliefs] and [Diversity] (the social dilemma context) model A is described. I only mention model B because it is more straightforward to treat mathematically. In [Behavior] and [Persistence] a third kind of noise is being used. Here an agent evaluate his expected payoff $U_i(t)$ in two cases: If he choose one action and if he choose the other. These two evaluations are then exposed to noise so that the results are normally distributed around the correct values. Then he compare the randomised values. All this would go on in the decision-boxes of my figures. I will not consider this kind of noise in this thesis, but if I did it would give the same results.

The new aspects of Huberman et al.'s approach to model social dilemmas are not that there are many agents, but the asynchronous dynamics, in continuous time is new, and the way they model the agents' expectations and beliefs about the future evolution of the game as we will look into now.

2.2 Expectations and strategies

We have now defined the exact rules of the game. Next we derive strategies of the players following [Beliefs] and [Collaboration].

An agent who updates at time t evaluates his expected payoff $U_i(t)$ in two cases: If he chooses C now, $U_i^C(t)$, and if he chooses D now, $U_i^D(t)$. He expects that $U_i^C(t)$ and $U_i^D(t)$ will be different because his choice now will affect the other agents in the future. He then chooses to cooperate if $U_i^C(t) > U_i^D(t)$ and to defect if $U_i^C(t) < U_i^D(t)$. He does not use a mixed strategy (with a randomised choice) except maybe if $U_i^C(t) = U_i^D(t)$ where he can choose with probabilities $(\frac{1}{2}, \frac{1}{2})$. In order to evaluate $U_i^C(t)$ and $U_i^D(t)$

an agent must have an internal model of how he expects the future dynamics to evolve. We saw in eq. 2.3 that an agent only need to have expectations about the future evolution of $\langle \hat{n}_c \rangle$ and $\langle \hat{k}_i \rangle$ in order to evaluate $U_i(t)$:

$$U_i(t) = \int_t^\infty (\frac{b}{N} \langle \hat{n}_c(t') \rangle - c \langle \hat{k}_i(t') \rangle) e^{-\frac{(t'-t)}{H}} dt' \tag{2.7}$$

Let $\hat{n}_c^C(t')$ and $\hat{k}_i^C(t')$ be the expected evolutions of $\langle \hat{n}_c(t') \rangle$ and $\langle \hat{k}_i(t') \rangle$ if he decides to cooperate and define $\hat{n}_c^D(t')$ and $\hat{k}_i^D(t')$ likewise. He decides to cooperate if and only if

$$\begin{aligned} 0 &< \Delta U_i(t) \equiv U_i^C(t) - U_i^D(t) \\ &= \int_t^\infty [\frac{b}{N} \hat{n}_c^C(t') - c\hat{k}_i^C(t') - \frac{b}{N} \hat{n}_c^D(t') + c\hat{k}_i^D(t')] e^{-\frac{(t'-t)}{H}} dt' \\ &= \int_t^\infty [b(\hat{f}_c^C(t') - \hat{f}_c^D(t')) - c(\hat{k}_i^C(t') - \hat{k}_i^D(t'))] e^{-\frac{(t'-t)}{H}} dt' \end{aligned} \tag{2.8}$$

$\Delta U_i(t)$ is the expected difference in utility. Let $\Delta \hat{f}_c(t') = \hat{f}_c^C(t') - \hat{f}_c^D(t')$ be the expected future difference between the fraction of agents cooperating if he chooses C now and if he chooses D. The agent expects that he will not change his own action in the future so that $\hat{k}_i^C(t') \equiv 1$ and $\hat{k}_i^D(t') \equiv 0$.[1] Then he decides to cooperate if and only if

$$\int_t^\infty [b\Delta \hat{f}_c(t') - c] e^{-\frac{(t'-t)}{H}} dt' > 0 \tag{2.9}$$

The central idea of their expectations is that *their action at the present will encourage similar actions in the future.* The reason may be imitation, establishing norms/conventions or that they expect others to have the same expectations and make the same decisions. So an agent believes that $\hat{f}_c^C$ relaxes exponentially to 1 and $\hat{f}_c^C$ relaxes exponentially to 0.

$$\begin{aligned} \hat{f}_c^C(t') &= 1 - (1 - \hat{f}_c^C(t)) e^{-\lambda(t'-t)} \\ \hat{f}_c^D(t') &= \hat{f}_c^D(t) e^{-\lambda(t'-t)} \end{aligned} \tag{2.10}$$

where λ is a parameter that controls the speed of this relaxation. He believes of course that his own choice at time t will make an instantaneous difference $\Delta \hat{f}_c(t) = 1/N$. See fig. 2.3.

[1] If there is noise (uncertainty) then $\hat{k}_i^C(t')$ and $\hat{k}_i^D(t')$ must move away from 1 and 0, so we make an error. But the horizon H reflects this uncertainty, and the discounting makes the error small.

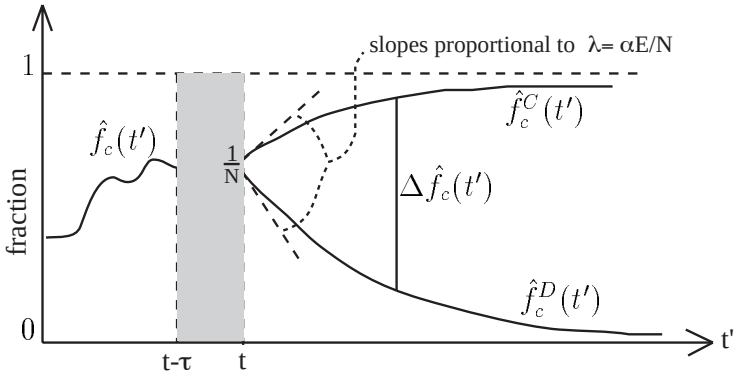


Figure 2.3: Anticipated evolution of the fraction of cooperators if C is chosen, and if D is chosen, at time t.

Since $\hat{n}_c$ can not decrease faster than when a cooperator turns to defection as soon as he is updated, we have

$$-\hat{n}_c\alpha \leq \frac{d\hat{n}_c}{dt} \tag{2.11}$$

Similarly there is an upper limit

$$\frac{d\hat{n}_c}{dt} \leq (N-\hat{n}_c)\alpha \tag{2.12}$$

so

$$-\hat{f}_c\alpha \leq \frac{d\hat{f}_c}{dt} \leq (1-\hat{f}_c)\alpha \tag{2.13}$$

From eq. 2.10 these limits gives us a bound on λ: $\lambda < \alpha$. λ is the rate at which the agent's action encourages similar action. It is both the rate of encouragement in the group of defectors and in the group of cooperators. It is thus a measure of the strength of an agent's influence. An agent believes his influence decreases with increasing group size N, because his decision makes a difference of b/N in others received utility. There are N agents who must share the influence. So we say $\lambda < \alpha/N$. Actually we will write

$$\lambda = \frac{\alpha E(\hat{f}_c(t-\tau))}{N} \tag{2.14}$$

i.e. we 'normalize' with the time scale set by α. The function E, with values less than one, is how strong an agent expects his influence to be. Sometimes we call this function the expectations of the agent. He expects this to depend on the present fraction of cooperators, which he takes the perceived fraction $\hat{f}_c(t-\tau)$ (with or without noise: model A and B) to be an estimator of.

We will at this point try to maintain generality, so we will not assume a specific functional form of E. Let us write the future difference $\Delta\hat{f}_c(t')$, that we need to do the integral in the condition eq. 2.9, that C is considered the best choice, as

$$\Delta\hat{f}_c(t') = 1 - (1 - \frac{1}{N})e^{-\frac{\alpha E(\hat{f}_c(t-\tau))}{N}(t'-t)} \tag{2.15}$$

Then the integral becomes

$$\begin{aligned} \Delta U_i(t) &= \int_t^\infty [b(1-(1-1/N)e^{-\frac{\alpha E(\hat{f}_c(t-\tau))}{N}(t'-t)}) - c]e^{-\frac{(t'-t)}{H}}dt' \\ &= H(b-c) - \frac{bH(N-1)}{\alpha E(\hat{f}_c(t-\tau))H + N}, \end{aligned} \tag{2.16}$$

and an agent chooses C if and only if

$$E(\hat{f}_c(t-\tau)) > \frac{1}{H\alpha}(\frac{Nc-b}{b-c}) \equiv E_{crit} \tag{2.17}$$

i.e. if his influence is strong enough!

2.2.1 'Bandwagon' and 'opportunistic'

We want to be able to capture different kinds of beliefs, which would depend on the social situation in real life, with different E functions. Also we should not talk about the exact form of E functions, only their rough shape. We should talk about classes of E functions. In [Beliefs] the five simplest imaginable classes of E functions are considered. Only two of them are considered important and relevant: 'bandwagon' expectations and 'opportunistic' expectations, represented by the functions $E(f_c) \sim f_c$ and $E(f_c) \sim 4f_c(1-f_c)$ respectively.

Bandwagon[2] expectations have an influence that is large when the group of cooperators is large and small when the fraction is small. An agent with

[2] To 'jump the bandwagon' means to join a successful movement.

bandwagon expectations believes cooperators to be much more sensitive to his action than defectors (who are sluggards)[3] so that the larger f_c the bigger sensitivity there is. To justify this belief as realistic I quote some examples.

> Consider bandwagon expectations in the context of working late in the office. If everyone on the team usually goes home early, then someone who stays longer will not expect others to follow her example – on the contrary, that person may be seen as an overly ambitious outsider. Once, however, a working atmosphere arises where people often work late, then peer pressure often leads individuals to go along. Moreover, defection by one team member is widely visible and will either be punished by the others or have a significant demotivating effect, leading to spreading defection by colleagues.—[Collaboration]

> Consider the set of beliefs the agent expects of others in the context of recycling programs. Recycling has a strong public good component because its benefits are available to all regardless of participation. Not too long ago very few towns had such programs. Perhaps you would read in the paper that a small town in Oregon had started a recycling program. Big deal. But several years later, when you read that cities all over your state have jumped onto the recycling bandwagon, then suddenly the long-term benefits of recycling seem more visible: recycled products proliferate in the stores, companies turn green, etc. Alternatively, imagine some futuristic time when everyone recycles, in fact your town has been recycling for years, everything from cans to newspapers to plastic milk jugs. Then you hear that some places are cutting back their recycling efforts because of the expense and because they now believe that the programs don't do that much

[3]There is a problem with this interpretation. If we have eq.2.10 then λ and then also the E function is the strength of influence on *both* the group of cooperators and the group of defectors. The defector group is thus not more sluggish in its response. However it should be possible to remedy this: What we only want and care about (and need in the integral eq.2.9) is the functional form eq.2.15. One can simply find an alternative to eq.2.10 that reflects the assumed difference in sensitivity in the two groups, while still leading to eq.2.15. To do this it would probably be best to reformulate exactly what model an agent forms of how another agent behaves.

> good after all. You think about all your wasted effort and imagine that the other towns still recycling are reaching the same conclusion. In view of this trend, your commitment to recycling may falter.—[Training, Beliefs]

The other class of expectations—opportunistic— have an influence which is increasing for increasing f_c when f_c is small, like bandwagon expectations have, but which decreases for large increasing f_c. With opportunistic expectations the influence is believed to be small in a mostly cooperating group. The belief is that a small amount of free riding will go unnoticed. One could for example believe that the cooperative effort have enough support so that additional contributions matter less. "They will probably manage without my help!". Opportunistic expectations were introduced to capture situations where the agents defect if f_c is large enough. Maybe because the cooperating agents accept a small fraction of defectors since a retaliation would also hit themself. Opportunistic expectations gives an interesting dynamics different from bandwagon. It is a bit more complex. It is considered in [Beliefs].

2.2.2 Simple rule

I will from now on assume bandwagon expectations $E(f_c) = f_c$. Then eq. 2.17 becomes

$$\hat{f}_c(t-\tau) > \frac{1}{H\alpha}(\frac{Nc-b}{b-c}) \equiv f_{crit} \tag{2.18}$$

What we have derived is an agent's strategy. It is the interior of the decision-boxes in model A and B. It turned out to be a very simple 'threshold rule' or 'barrier rule' or 'majority rule':

$$k_i(t) = \Theta(f_c^i(t-\tau) - f_{crit}) \tag{2.19}$$

where Θ is the Heaviside function. It is exactly the functions in the bandwagon class that give such a rule.

When an agent updates he first gets the perceived fraction of cooperators from his input. This may be noisy (model B) or not (model A). If the fraction is greater than f_{crit} he intends to cooperate, otherwise defect. His intention is carried out with (A) or without (B) noise, and his new state is obtained. This is very suited for Monte Carlo simulations (which is no surprise).

In the study of multiagent systems and distributed artificial intelligence one can distinguish between *cognitive* and *reactive* agents. Cognitive agents are implemented with a sophisticated high level internal representation of the outer world and plans for the future. Reactive agents do not have have a complex internal structure. Their behaviour is given as a simple function of the input. Interestingly, we began discussing 'cognitive' elements like expectations and beliefs but ended up with simple reactive agents. From our long discussion only the constant f_{crit} survives into our simulations.

2.3 Dynamics

In the last section the stochastic behaviour of an individual agent was fully specified. What we are interested in, though, is the collective behaviour af many agents. The next task is therefore to investigate the global dynamics. This task will lead us into a new domain of analysis and we can leave behind us utility functions and expectations. First we will see if we can describe the dynamics as a Markov process. We shall see that we can do that, and that we can map the dynamics onto the Glauber dynamics of an Ising model. But this is only possible for model B. I will refer to A.2 and A.3 for a presription of how to construct a Markov process for the system.

2.3.1 Model B

First we consider model B. Let us try the system states $\hat{k}$ as our Markov states. The rate of updates of the system is $N\alpha$ where α is the decision rate of an agent. Since an agent may decide to do the same action as before, we can have 'no jump' (A.2). Assume the present state is $\hat{k}$ and that agent i updates. He receives information about $\hat{n}_c(t-\tau)$. But since jump probabilities should depend only on the present state and not the past we must have no delay $\tau = 0$. We will by ρ denote the probability that agent i chooses C given the state $\hat{k}$. There is no noise on his output so this is also the probability that C is evaluated to be the best action plus half the probability that C and D are considered equally good.

$$\rho \;=\; P(f_c^i > f_{crit}) + \frac{1}{2}P(f_c^i = f_{crit})$$

$$= \quad P(S_p + S_q > Nf_{crit}) + \frac{1}{2}P(S_p + S_q = Nf_{crit}) \tag{2.20}$$

where we assume $p = q$ and $S_p \in B(\hat{n}_c, p)$ and $S_q \in B(N - \hat{n}_c, 1 - p)$. This expression for ρ can be written out using the binomial distributions (see app. A in [Outbreak]) but we will use that a stochastic variable $X \in B(n, p)$ is approximately normally distributed $X \in N(np, np(1-p))$ when np and $n(1-p)$ both are greater than about 5. So we take

$$\begin{aligned} S_p &\in N(\hat{n}_c p, \hat{n}_c p(1-p)) \\ S_q &\in N((N - \hat{n}_c)(1-p), (N - \hat{n}_c)(1-p)p) \end{aligned} \tag{2.21}$$

n_c^i is the sum of S_p and S_q and

$$\begin{aligned} \langle S_p + S_q \rangle = \langle S_p \rangle + \langle S_q \rangle = \hat{n}_c p + (N - \hat{n}_c)(1-p) \\ \sigma^2(S_p + S_q) = \sigma^2(S_p) + \sigma^2(S_q) = N(1-p)p \end{aligned} \tag{2.22}$$

since S_p and S_q are independent. A sum of two independent normally distributed variables is again a normally distributed variable. So

$$n_c^i \in N(\hat{n}_c p + (N - \hat{n}_c)(1-p), N(1-p)p) \tag{2.23}$$

and

$$f_c^i = \frac{n_c^i}{N} \in N\left(\hat{f}_c p + (1 - \hat{f}_c)(1-p), \frac{(1-p)p}{N}\right) = N(\mu, \sigma^2) \tag{2.24}$$

where $\mu = \hat{f}_c p + (1 - \hat{f}_c)(1-p)$ and $\sigma^2 = (1-p)p/N$. Then ρ becomes

$$\rho = P(f_c^i > f_{crit}) = 1 - \Phi\left(\frac{f_{crit} - \mu}{\sigma}\right) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\mu - f_{crit}}{\sigma\sqrt{2}}\right) \tag{2.25}$$

where $\Phi(x)$ is the area under the standard normal curve from $-\infty$ to x, and we use the error function $\mathrm{erf}(x) = 2\Phi(x\sqrt{2}) - 1$. ρ is a function of $\hat{f}_c$ only (through μ). When we know $\rho(\hat{f}_c)$ we can find all the jump probabilities we like. They will depend only on the present state of the system and we have specified a Markov process. We shall see that eq. 2.25 is similar to the corresponding equation for an Ising model.

2.3.2 Model A

Now consider model A. Again we must have $\tau = 0$ and the decision rates are as before. First we try having $\hat{k}$ as our Markov states as before. Assume the present state is $\hat{k}$ and that agent i updates. The agent knows the present $\hat{f}_c$ and

$$\rho = P(\hat{k}_i = 1) = \begin{cases} p & for \quad \hat{f}_c > f_{crit} \\ 1-q & for \quad \hat{f}_c < f_{crit} \end{cases} \tag{2.26}$$

This will not look like Glauber dynamics. Let us instead have the *intentions* k as our Markov states. Assume the present state is k and that agent i updates. He receives the information $\hat{f}_c$ and

$$\rho = P(k_i = 1) = P(\hat{f}_c > f_{crit}) + \frac{1}{2}P(\hat{f}_c = f_{crit}) \tag{2.27}$$

And given k these probabilities are

$$\rho = P\left(\frac{S_p + S_q}{N} > f_{crit}\right) + \frac{1}{2}P\left(\frac{S_p + S_q}{N} = f_{crit}\right) \tag{2.28}$$

where $p = q$ and $S_p \in B(n_c, p)$ and $S_q \in B(N - n_c, 1 - p)$. Like for model B we get the approximation

$$\rho(f_c) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\mu - f_{crit}}{\sigma\sqrt{2}}\right) \tag{2.29}$$

where $\mu = f_c p + (1-f_c)(1-p)$ and $\sigma^2 = (1-p)p/N$. Again jump probabilities can be found. However they will not only depend on the state k. They also depend on the actual state $\hat{k}$. So decisions are not independent. Succesive decisions are correlated since $\hat{k}$ is either unchanged or changed only by ± 1. If we want a Markov process to describe the evolution of k we must use $(k, \hat{k})$ as our state. On the other hand the probabilities for the next $\hat{k}$ do not depend on k. I do not know how the correlations from the underlying extra memory of $\hat{k}$ affects the dynamics of the intentions k and the attempted level f_c of cooperation, so I prefer model B. Maybe there is no difference in the long run.

2.4 Mapping onto Ising model

Consider an N-spin infinite range/mean field Ising model with Glauber dynamics (see app. B). Each spin is updated with a rate α. The Hamiltonian

is

$$H = -\frac{1}{2}\sum_{i,j}\frac{1}{N}s_i s_j - h_0\sum_i s_i \tag{2.30}$$

where $\sum_{i,j} = \sum_i \sum_{j\neq i}$. The effective field on spin i is

$$h_i = h_0 + \sum_{j\neq i}\frac{1}{N}s_j = h_0 + m = h_0 + 2\frac{n}{N} - 1 = h_0 + 2f - 1 \tag{2.31}$$

where m is the magnetisation, n is the number and f the fraction of spins equal to $+1$. The probability that s_i is equal to $+1$ after it has been updated is (see eq. B.8)

$$\begin{aligned}\rho_{+1} &= P_{accept}(-2h_i) \\ &= \frac{1}{2} + \frac{1}{2}\tanh\left\{(2\beta)f + \beta(h_0 - 1)\right\}\end{aligned} \tag{2.32}$$

Compare with eq. 2.25:

$$\rho = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left\{(\frac{2p-1}{\sigma\sqrt{2}})\hat{f}_c + (\frac{1-p-f_{crit}}{\sigma\sqrt{2}})\right\} \tag{2.33}$$

If we use the approximation $\mathrm{erf}(x) \approx \tanh(x)$ then the two models' dynamics are identical if the two linear expressions inside the brackets are identical, i.e.:

$$2\beta = \frac{2p-1}{\sigma\sqrt{2}} \tag{2.34}$$

$$\beta(h_0 - 1) = \frac{1-p-f_{crit}}{\sigma\sqrt{2}} \tag{2.35}$$

where $\sigma = \sqrt{p(1-p)/N}$. This gives us

$$\beta = \frac{(2p-1)\sqrt{N}}{\sqrt{8p(1-p)}} \tag{2.36}$$

$$p = \frac{1}{2} + \sqrt{\frac{\beta^2}{2N + 4\beta^2}} \tag{2.37}$$

(this relation between β and p is shown in fig. 2.4) and

$$h_0 = \frac{1 - 2f_{crit}}{2p-1} \tag{2.38}$$

$$f_{crit} = \frac{1}{2} - \frac{\beta h_0}{\sqrt{2N + 4\beta^2}} \tag{2.39}$$

These four equations provide a mapping back and forth between the social

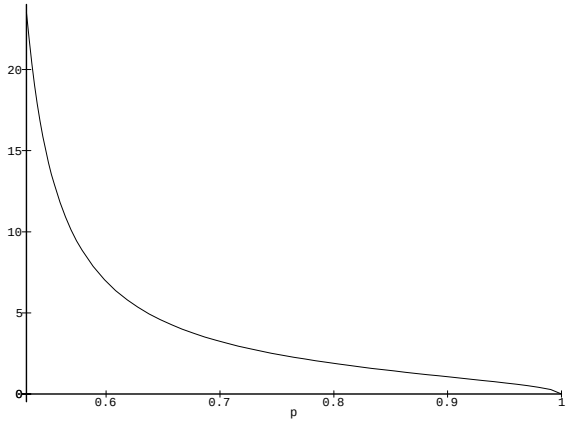


Figure 2.4: $TN^{1/2}$ as a function of p. The curve has vertical slope at $p = 0.5$ and $p = 1$.

model and the Ising model. The dynamics of the social model is simply the Glauber dynamics of the corresponding Ising model with tanh replaced by erf.

It is strange that we get a symmetry between cooperation and defection like the symmetry between positive and negative magnetization in the Ising model. There is not a symmetry between C and D in the utility function! It has to do with the symmetric way we have noise ($p = q$).

2.4.1 Global properties

I will now apply our knowledge of the Ising model to the social model. With infinite range interactions mean field theory (app. B) is very accurate (it becomes exact when $N \to \infty$). According to it the average magnetizations m of the equilibrium states are given as solutions to

$$m = \tanh[\beta m + \beta h_0] \tag{2.40}$$

m is also the mean $\langle s_i \rangle$ for each i. We are only interested in states (fixed points) that are stable, i.e. minima of the free energy function (see fig. 2.5). The critical temperature is $\beta_c = 1/T_c = 1$. For $\beta \leq \beta_c$ there is only one

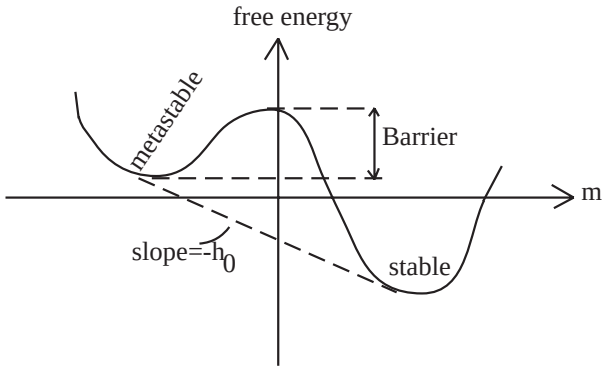


Figure 2.5: Free energy picture

stable equilibrium state. The sign of m in that state is equal to the sign of h_0. If $h_0 = 0$ then $m = 0$. For $\beta > \beta_c$ there are two stable states, one with positive m and one with negative m.

A positive m corresponds to a mostly cooperating state in the social model and a negative m is a mostly defecting state. The critical 'temperature' or noise level is (with $\beta = 1$ in eq. 2.37)

$$p_c = \frac{1}{2} + \frac{1}{\sqrt{2N+4}} \tag{2.41}$$

or

$$\sigma_c = \frac{1}{\sqrt{4N+8}} \tag{2.42}$$

So, for $p > p_c$ there are two possible states. They are given as minima of the free energy function shown in fig. 2.5. The deepest (global) minimum is an optimally stable state. It has the sign of h_o. The other state is metastable. Fluctuations will eventually take the system over the free energy barrier to the optimal stable state in a time which is exponential in the barrier height.

At $h_0 = 0$ both states are optimal and coexisting. Here fluctuations will take the system from one to the other (flip-flop). This special value $h_0 = 0$ corresponds, when inserted in eq. 2.38, to $f_{crit} = 1/2 \equiv \tilde{f}^*_{crit}$. So for $f_{crit} < 1/2$ a mostly cooperating state is optimal and for $f_{crit} > 1/2$ a mostly defecting state is optimal. When $p < p_c$ and $f_{crit} = 1/2$ there is one state, and it has equally many cooperators and defectors.

Actually, when $p > p_c$, the metastable state disappears if $|h_0|$ is large enough. As h_0 increases in fig. 2.5 the barrier is lowered until at $h_0 = \tilde{h}$ it becomes zero and the metastable fixpoint with $m < 0$ becomes marginally stable. With opposite signs the same holds: At $h_0 = -\tilde{h} \equiv h^*$ the metastable state with $m > 0$ becomes marginally stable[4]. For $|h_0| > \tilde{h}$ there is no metastable minimum, only a global maximum. The critical value $\tilde{h}$ is a function of β and can in principle be found using the condition that eq. 2.40 has two and not three solutions, see fig. 2.6.

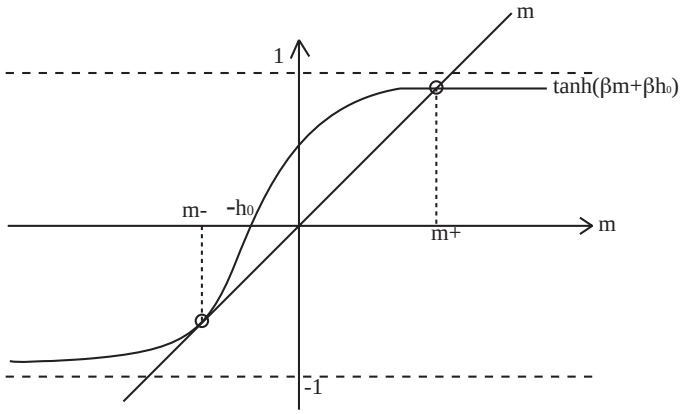


Figure 2.6: Here $h_0 = \tilde{h}$ so there are two solutions, $m-$ and $m+$, to $m = \tanh[\beta m + \beta h_0]$. If h_0 is increased then $m-$ will disappear and if h_0 is decreased the equation will have three solutions.

These results are summarized in the phase diagram fig. 2.7 And using the mapping equations 2.37 and 2.38 we can in fig. 2.8 show the same phase diagram with f_{crit} and p instead of h_0 and T.

The critical value $\tilde{f}_{crit}$ (and its counterpart $f^*_{crit} = 1 - \tilde{f}_{crit}$) can be found when there is no noise, $p = 1$: Then an agent will defect if $f^i_c = \hat{f}_c < f_{crit}$. A mostly defecting metastable state is no longer possible if $f_{crit} < 0$ since the smallest value $\hat{f}_c = \sum_{j=1}^{N} \hat{k}_j/N$ can take is 0. So $\tilde{f}_{crit}(p = 1) = 0$. And then $f^*_{crit}(p = 1) = 1$. And this is because the largest value $\hat{f}_c$ can attain is 1. Mapping these values back to h_0 we get $\tilde{h}(T = 0) = 1$ and $h^*(T = 0) = -1$.

[4]When this is the case we are at the 'spinodal'.

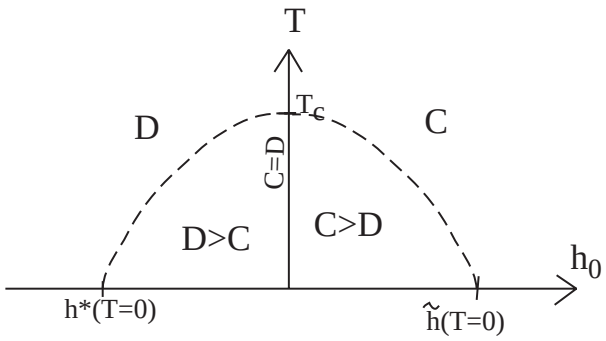


Figure 2.7: $T - h_0$ phase diagram. Letters C and D indicates possible states and the notation $D > C$ means D-state is more stable than C-state (which is metastable). $C = D$ means that there are two equally stable states. The boundaries where the metastable state disappears (given by functions $\tilde{h}$ and h^* of T) are sketched with dashed lines, since I have not calculated the exact shapes.

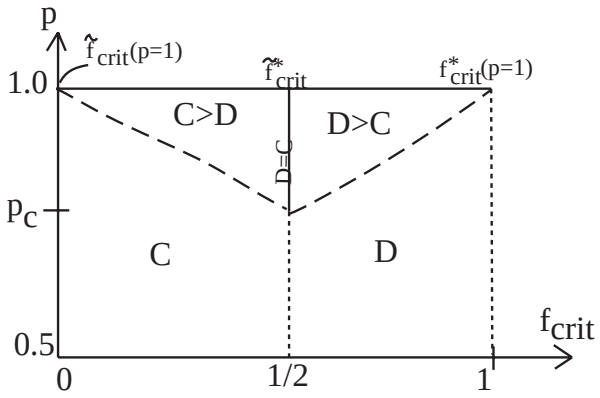


Figure 2.8: $p - f_{crit}$ phase diagram. Letters C and D indicates possible states and the notation $D > C$ means D-state is more stable than C-state (which is metastable). $C = D$ means that there are two equally stable states. The boundaries where the metastable state disappears (given by functions $\tilde{f}_{crit}$ and f^*_{crit} of p) are sketched with dashed lines, since I have not calculated the exact shape.

However, for the Ising model $\tilde{h}(T=0)$ is exactly h_{max} defined in app. B.

$$\tilde{h}(T=0) = h_{max} = \max_i \sum_{j \neq i} J_{ij} = \frac{N-1}{N} \tag{2.43}$$

which is not equal to 1. The reason is that an agent also counts himself while an Ising spin does not interact with itself.

Eq. 2.18 gives f_{crit} as a function of group size N and horizon H. From that we get

$$N = \frac{b}{c} + \alpha H f_{crit}(\frac{b}{c} - 1) \tag{2.44}$$

which we use to translate the tilde and star quantities and make a $N - H$ phase diagram for some $p > p_c$ (i.e. $T < T_c$) in fig. 2.9

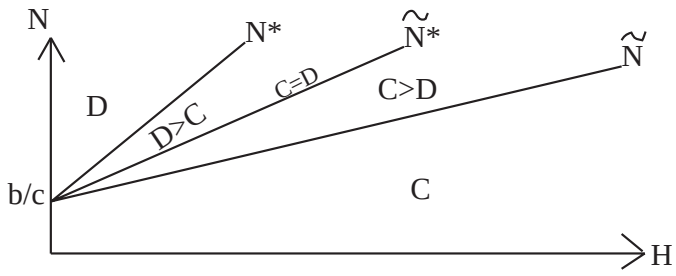


Figure 2.9: $N - H$ Phase diagram for a temperature below the critical. The three boundaries between domains are given by the functions N^*, $\tilde{N}^*$ and $\tilde{N}$ of H.

2.4.2 Interpretation

With fig. 2.9 we can try to understand our results in the social dilemma context: If the group size N drops below b/c only one equilibrium is possible: a mostly cooperating one. This can be interpreted via the utility function eq. 1.1. $N > b/c$ was one of the conditions for a social dilemma. If $N \leq b/c$ a player will always get more in return (the marginal benefit) when he chooses C than what he gives (the cost), so C is always the best choice.

We see in the diagram that cooperation is sustainable only if the horizon H is big enough. This was also the result with the repeated game in discrete

time (the folk theorem), where the discount parameter had to be big enough. It is also common sense that cooperation requires a long term interaction. Furthermore, the stability of the cooperative phase relative to the defective phase increases with increasing horizon.

Finally we see there is an upper limit N^* on the group size above which cooperation is not sustainable. This is in accordance with notions in sociology [Olson, Bendor, Ostrom]. And with increasing group size there is a decreasing stability of the cooperative phase relative to the defective phase. This result has root in our assumption that an agent believes his influence to be inversely proportional to N, since he only chooses C if his believed influence is big enough.

These three results are not the most significant of our model. They are more or less put in to begin with. Most interesting is the picture that our model gives of the dynamics of social dilemmas. According to this picture the social system can sit in (meta)stable states, with small fluctuations, for long times until suddenly the fluctuations changes the overall configuration.

2.4.3 Alternative derivations

I noticed that tanh $\approx$ erf and assumed then that the global properties of the social and Ising models are the same, but in general one should be cautious when going from local to global like that (cf. chaos). Huberman *et al.* derive the global properties correctly using the error function. They use an Ω-formalism, where Ω is an optimality function or free energy. I could have compared the Ω of the social model with the free energy or Ω of the Ising model and thereby made a mapping on a more solid basis. But that is not necessary: Qualitatively they yield the same results and phase diagrams. Only the function $f^*_{crit}(p)$, which I did not calculate (dashed lines), would probably depend a little on the difference between tanh and erf, but only for $p < 1$: Like me they get $f^*_{crit}(p = 1) = 1$ and the corresponding upper critical group size

$$N^*(H) = \frac{b}{c} + \alpha H(\frac{b}{c} - 1) \tag{2.45}$$

They also get for all p that $\tilde{f}^*_{crit}(p) = 1/2$ and

$$\tilde{N}^*(H) = \frac{b}{c} + \frac{1}{2}\alpha H(\frac{b}{c} - 1) \tag{2.46}$$

But in [Outbreak] they calculate as if the lowest value the perceived $\hat{f}_c$ can attain is $1/N$ while the highest is 1. So they get $\tilde{f}_{crit}(p=1) = 1/N$ and

$$\tilde{N}(H) = \frac{b}{2c} + \frac{1}{2c}\sqrt{b^2 + 4H\alpha c(b-c)} \tag{2.47}$$

for $p = 1$, while I get $\tilde{N}(H) = b/c$. Apparently they use that agent i perceives

$$\hat{f}_c = 1 - \frac{1}{N}\sum_{j\neq i}(1 - \hat{k}_j) \tag{2.48}$$

i.e. he do not count himself in the sum.

The Ω-formalism uses eq. 2.25 as starting point, like our analysis. But that expression is an approximation itself. Theory should therefore be compared with simulations, and Huberman *et al.* also do that, of course. In simulations it is also possible to have $p \neq q$ and $\tau > 0$.

2.4.4 Terminology

When we talk about the 'optimal' equilibrium state as opposed to the metastable state we do not mean 'optimal' in terms of utility to the agents or an optimal level of cooperation in terms of effectiveness or productivity. And we do not talk about optimal strategies. The word optimal only refer to the global minimum of the free energy. And when we talk about 'equilibria' we do *not* mean Nash equilibria (NE) or evolutionarily stable strategies abbreviated ESS (see [MaynardSmith]). The dynamics changes the actions of the agents and moves the system in *configuration space* until an equilibrium is approached. But a NE refers to a set of strategies so it is a point in *strategy-profile space*. A strategy of an agent is for example the threshold rule eq. 2.19. The set of strategies (rules) is fixed during the dynamics in configuration space—it determines the dynamics!— and can therefore not come to an equilibrium dynamically.

In [Outbreak, Diversity] an equilibrium of the dynamics is actuallly called a Nash equilibrium. They also refer to actions as 'strategies'. Actions and strategies may be identified if the strategy says: 'Play C!' in the one shot game or 'Play C always' in the ongoing game or 'Play C in x percent of the time'. But these are not the decision rules we have. And only one NE can

be made that way: Everybody always plays D. So the use of the words in the articles is not compatible with my use.

Our equilibria of the dynamics are called ESS (evolutionarily stable strategy) in [Persistence, Behavior]. But an ESS is a strategy of a single individual with the property that a population where everyone has adopted this strategy can not be invaded by a mutant. (This concept is close to a NE). Probably they mean an evolutionarily stable *state*, which is a stable distribution of genes (or actions?) in a population.

We must make a distinction between dynamics in configuration space and strategy dynamics. Actions are being changed. But our strategies are fixed before the game is played. They are not tested and then improved. There is no feedback to the strategies from the actual dynamics. This will be discussed in the next section.

2.5 Discussion

I will discuss how sensible the expectations eq. 2.10 are. First I see three minor problems:

- An agent expects his action to encourage similar actions. I find it hard to interpret this in other ways than that the action is a *signal* to the other agents that influences their choices. But in our agent model (A or B) the agent can only observe the fraction at time $t-\tau$ and not actions which are *changes* in this fraction.

- The effect of the action D depends on who takes it: Assume you are an agent who chooses D at time t. From your expectations eq. 2.10 we get

$$\frac{d\hat{f}_c^D}{dt'}(t') = -\frac{\alpha \hat{f}_c(t-\tau)}{N}\hat{f}_c^D(t') \approx -\frac{\alpha}{N}\hat{f}_c^D(t)\hat{f}_c^D(t') \tag{2.49}$$

  where we assume $\hat{f}_c(t-\tau) \approx \hat{f}_c^D(t)$. The present fraction is $\hat{f}_c^D(t) = f_1$. At some later time $t' = t_2$ the fraction has dropped to $\hat{f}_c^D(t_2) = f_2 < f_1$. Another agent chooses D at time t_2 and it results in a rate of change equal to $-\frac{\alpha}{N} f_1 f_2$. But had you made your action when the present fraction $\hat{f}_c^D(t)$ was f_2 instead of f_1, then your action would have resulted in a rate of change equal to $-\frac{\alpha}{N} f_2^2$ which is less than what the other

agent accomplishes. So, at the same level of cooperation an action has an effect depending on who takes it.

- The noise $1-p$ is not taken into account by the agents, in their considerations leading to the rule eq. 2.19, because their expectations are independent on the noise. A possible way to include noise is to let the horizon H reflect noise. With a high noise level the horizon should be shorter, to reflect the higher uncertainty about the future. Then H and p will not be independent variables.

These problems are a matter of interpretation. But there is also a major problem:

- The expectations are not consistent with the actual dynamics. The time evolution of $\langle \hat{f}_c(t') \rangle$ is given by the equation of motion (see eq. 6 in [Behavior])

$$\frac{d\langle \hat{f}_c \rangle}{dt'} = \alpha(\langle \rho \rangle - \langle \hat{f}_c \rangle) \tag{2.50}$$

  Here we can use the mean field approximation $\langle \rho \rangle = \rho(\langle \hat{f}_c \rangle)$ where $\rho(\hat{f}_c)$ is given by eq. 2.33. Below the critical temperature we have f_-, f_0, f_+ as the three mean field solutions to $\langle \hat{f}_c \rangle = \rho(\langle \hat{f}_c \rangle)$. When $\langle \hat{f}_c \rangle > f_0$ then $\langle \hat{f}_c \rangle \to f_+$ and when $\langle \hat{f}_c \rangle < f_0$ then $\langle \hat{f}_c \rangle \to f_-$, regardless of what action the agent chooses. Only when $\langle \hat{f}_c \rangle \approx f_0$ we can have that the basin of attraction is determined by the action.

  What happens if an agent uses this behaviour of $\langle \hat{f}_c \rangle$ as expectations instead of eq. 2.10 ? If the fraction is large enough the agent only expects he has a small influence on the speed at which $\langle \hat{f}_c \rangle$ settles in f_+. So then it would probably pay to choose D. Then perhaps we would get a rule like the one for opportunistic expectations. Or perhaps we would get that D always is the best choice. But if the agents no longer use the rule eq. 2.19 then eq. 2.33 no longer holds. What we therefore could be looking for is a rule and some expectations, where the expectations would be consistent with eq. 2.50 for that rule, and where the rule is rational given the expectations. This idea is the same as the one behind the game theoretic solution concept called a *sequential equilibrium*. Rational agents should behave like that.

These problems present a criticism to the model. It attacks the imperfectness of the agents and the imperfectness of the mathematical derivations.

The answer to the criticism concerning imperfectness of the agents is [Training, Collaboration] that agents are not 100% rational. Expectations are only beliefs and not perfect knowledge. Agents do not carry out a game theoretic analysis. We are allowed to assume a system of agents with some set of expectations, even if they are irrational.

In many systems it is natural that strategies evolve (sect. 1.3.2). Agents do not have to be rational but they can learn about the dynamics and then discover better strategies or better expectations. The evolution process where strategies are chosen takes place on a larger time scale than the game dynamics where actions are chosen. If we look at some stage of the evolutionary dynamics where it has not come to an equilibrium then the agents' expectations and strategies may very well have inconsistencies.

Of course more perfect agents are also interesting. What happens if the agents' models and expectations about the future are at least qualitatively correct? This is a possible area for further work on the flat model[5].

For an answer to the criticism concerning the imperfectness of the mathematical derivations I will refer to a general discussion of this in the conclusion at the end of this thesis (section 8.1).

My supervisors advised me to accept the model and go on with the *hierarchical model* instead, which we will see in the next chapter. A threshold rule like eq. 2.19 intuitively seems to be a good strategy. This intuition is also confirmed by the succes of another threshold rule namely TIT FOR TAT. An intuitively justified rule is relevant since people behave intuitively more often than following mathematical reasoning. The model also gets many things right. We also expect threshold rules to underlie the dynamics of social systems in many contexts different from ours. Therefore, the model is qualitatively interesting in a broader context.

[5] With simulations of evolution/learning or analytically one can look for invading strategies that can exploit a population using the rule eq. 2.19. Then one can look for a NE or ESS or some expectations that are consistent with the resulting strategy. It would be interesting, for example, to find that a second social dilemma appeared on strategy level rather than action level. A strategy with a higher tendency to choose D would perhaps give a short term increase in payoff but a long term loss, because the other individuals would also turn to a strategy like that. That requires a long horizon of course.

Chapter 3

The hierarchical model

In this chapter I describe the *social hierarchical model* with [Fluidity] as the main reference. [Fluidity] deals with fluidity—changing organizational structures—but I will only look at fixed hierarchical structures. The derivation of the model in this chapter will be less explicit and rigorous than the derivation of the flat model in the previous chapter.

3.1 Social structures

In the flat model in the previous chapter there was no structure in the group of individuals. Everybody interacted equally with each other. In sociology one studies groups that *have* a social structure with different kinds of relations between people [Network, Burt]. These different relations may for example be affective ties or how often they interact. Maybe we can describe this with a matrix A where A_{ij} is the link or the interaction strength between individuals i and j. A defines a network of links. Maybe this social network of relations has a topology-like structure so that individuals interact stronger the closer they are to each other. Sometimes our interactions are restricted physically by the spatial world we live in (geographically). But of course people who are neighbors may not have any interaction at all, and distances are shortened with communication technology.

Varying strengths of interactions can have importance in social dilemmas. A good example is air pollution. People who limit the pollution from their chimney will give a greater benefit to people who live nearby than to people

living far away. A similar example is slowing down arms races. Maybe the problem of the many hand weapons in the U.S. can also be described this way. In these examples the metric is euclidian. This could suggest that we made a lattice Ising model, such as physicists are used to. But we will not do that, because people living on the surface of the earth are not uniformly distributed. They live in clusters for example towns, and towns may again form clusters in metropolitan areas. Clusters form clusters in a hierarchical way: Buildings, neighbourhoods, towns, states, continents, planets etc. Hierarchical structures also appear in examples of social dilemmas with no connection to geography. Societies and human organizations are often organized hierarchically. In fact hierarchical structures are quite general. So we will instead of spatial distance use the *ultrametric distance*.

3.2 The ultrametric distance

Imagine a group of agents, or just *sites*, which is divided into clusters or subgroups in a hierarchical way. This can be illustrated by a *tree* as in fig. 3.1 where the sites are the endpoints of the tree. The branching points

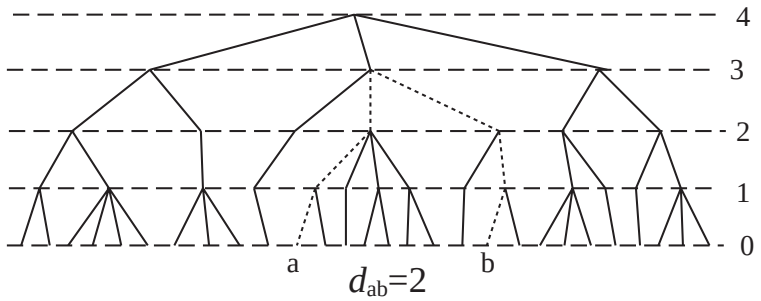


Figure 3.1: Hierarchical system and redefined ultrametric distance.

of the tree are called *ancestor nodes*. The ancestor nodes may branch into only one or more than one branch. The *levels* on which ancestor nodes reside are numbered $1, 2, 3, \ldots$ as in fig. 3.1. If L is the number of levels then the top ancestor node (the root) is on level L. The sites (endpoints) are all on level 0. Each site is the descendant of a unique ancestor node on each level. We define the *ultrametric distance* d_{ij}^{UM} between sites i and j to be the first level on which their ancestors coincide. So sites with a common ancestor on

level 1 have ultrametric (UM) distance 1, and sites with a common ancestor on level 2, but not on level 1, have UM distance 2 and so on. We define the distance between a site and itself to be 0. This is a metric because

- $d_{ij}^{UM} \geq 0,\ d_{ij}^{UM} = 0 \Leftrightarrow i = j$
- $d_{ij}^{UM} = d_{ji}^{UM}$
- $d_{ij}^{UM} = \max\{d_{ik}^{UM}, d_{kj}^{UM}\}$, so the triangular inequality holds.

For the sake of convenience we will redefine the distance: We will instead use $d_{ij} = d_{ij}^{UM} - 1$ and in the rest af the thesis this distance is called the ultrametric distance. But we still define $d_{ii} = 0$.

Ancestor nodes are *not* sites in the model. And the hierarchy does not imply that sites have different ranks. The sites in fig. 3.1 are ordered in a linear chain, with two neighbours and so on, but this ordering is not part of the system. No site is the center and there are no ends or boundaries. I define a cluster on the l'th level (a *l-level cluster*) to be the sites that are descendants of an ancestor node on level l. If the branching rate is constant and equal to z, i.e. every ancestor branches into z branches, then the number of sites is z^L, where L is the number of levels. In this case all sites are equivalent. Then I define a *block of z^l sites* or a *l-level block* to be the z^l descendants of a single ancestor node on level l.

3.3 The social hierarchical model

Imagine a social dilemma involving N agents who are hierarchically divided into smaller clusters. We have ultrametric distances in the system. As in the air pollution example agents only have little interaction with distant agents. They still interact with everybody, but mostly with members within the subgroup to which they belong. The effect of clustering is therefore to diminish the group size. This is the idea behind clustering. We saw in the previous chapter that reducing the group size means that the cooperative phase becomes more likely and stable.

Clustering also lowers the free energy barrier. In a ferromagnetic system it corresponds to short range interactions, while a flat model corresponds to infinite range interactions. And the nucleation barrier is smaller for short range interactions.

Clustering, or organizing the system in smaller units, is therefore proposed as a practical recipe for making organizations and corporations more cooperative [SciAm] and helping them out of metastable states, for example systems of computational agents [Persistence].

Clustering is also suggested to be a mechanism that helps the initiation of cooperation in biological evolution [Altruism, page 283].

We will try to extend the flat model to capture this idea. We will let A_{ij} describe the decaying interaction with UM distance. The matrix A could describe many kinds of social interactions, but here it should be their mutual influence in terms of utilities, since, in our approach, utilities determine their behaviours. (Maybe other social structures would have secondary effects). We will let A_{ij} decrease exponentially:

$$A_{ij} = a^{-d_{ij}} \tag{3.1}$$

and the contribution to the benefit of agent i from agent j (if j contributes) should be proportional to A_{ij}. Let the utility function of agent i be

$$u_i = b\frac{\bar{n}_{ci}}{\bar{N}_i} - c\hat{k}_i \tag{3.2}$$

where as usual $\hat{k}_i$ indicates whether or not i cooperates, and

$$\bar{n}_{ci} = \sum_{j=1}^{N} \hat{k}_j A_{ij} \tag{3.3}$$

is the *effective* number of cooperators from i's point of view, and

$$\bar{N}_i = \sum_{j=1}^{N} A_{ij} \tag{3.4}$$

is the *effective* group size from i's point of view. The flat model has $A_{ij} = 1$ for all i, j and corresponds to the special case where everybody belongs to the same cluster on the lowest level. The effective group size is in general smaller than the actual group size: $\bar{N}_i < N$. If everybody cooperates then the utility to an agent is $b - c$ and it should be positive if there is a social dilemma. Defection should also be the best one-shot action so for all i

$$b > c > \frac{b}{\bar{N}_i} \tag{3.5}$$

We define the effective fraction of cooperators from i's point of view:

$$\bar{f}_{ci} = \frac{\bar{n}_{ci}}{\bar{N}_i} \tag{3.6}$$

which is between 0 and 1.

It is possible to carry out the arguments and calculations of section 2.2 with $\hat{f}_c$ replaced by $\bar{f}_{ci}$ and N replaced by $\bar{N}_i$. An agent believes he can encourage similar actions in the group with effective size $\bar{N}_i$ where his influence is bigger. We get the condition that agent i cooperates if the percieved $\bar{f}_{ci}$ is greater than

$$\bar{f}_{crit,i} \equiv \frac{1}{H\alpha}\left(\frac{c\bar{N}_i - b}{b-c}\right) \tag{3.7}$$

The smaller effective group size thus means a smaller threshold compared to eq. 2.18. The threshold corresponded to an external field, and we see that the effect of clustering is indeed to increase the field in the direction of the mostly cooperating equilibrium. But we see that the external field $\bar{f}_{crit,i}$ depends on the site i.

3.4 Simulations

One can have noise as in model A or model B and then run simulations. Simulations of course show that cooperation is a more likely outcome than with a flat structure.

Simulations also reveal a *cascading phenomenon*. This can occur when the system goes from a metastable state of all defectors to the state of all cooperators (or the other way, depending on $\bar{f}_{crit,i}$). Cooperation first spreads within a 1-level cluster and after a while everybody in that cluster cooperates while the rest defects. Cooperation then spreads to the rest of the 2-level cluster which the first 1-level cluster is part of. Cooperation spreads in that way to clusters on successive levels. The turning to cooperation in a cluster *triggers* cooperation within the cluster on the next higher level.

Situations occur where a cluster on some level is filled with cooperators while the rest of the system defects, and such states may live for long times. They are metastable. The cooperators see a fraction above their thresholds while the rest see a fraction below their thresholds.

If we plot the number of cooperators versus time in such a run then the curve typically looks like a staircase, where the flat parts correspond to filled metastable clusters.

We would like to understand this cascading phenomenon. Is it a nucleation process? Can we describe it metaphorically like a step ladder: The metastable states are steps on the ladder and by using a ladder the system can move much faster than if the transition to cooperation were taken in one giant step? A difficult job is split up into smaller easier jobs? These questions will be answered in the next chapter.

3.5 Connection to Dyson's hierarchical model

We saw in section 2.4 how the flat social model could be mapped onto an Ising spin model with Glauber dynamics. That is also possible for a hierarchical model with a constant branching rate equal to z (this means that the external field becomes uniform). There is a spin model with interactions that has a hierarchical structure: *Dyson's hierarchical model* (see the next chapter). With Glauber dynamics and $\lambda = 1$ in Dyson's hierarchical model it corresponds to $a = z$ in the social hierarchical model. To make the mapping it is necessary to find the mean and variance of $\bar{f}_{ci}$ (we assume it is a Gaussian). This can be done with the aid of some mean field assumptions. But we will not do that. In the rest of the thesis we will be interested in the dynamics of Dyson's hierarchical model, knowing that it maps onto the social hierarchical model.

Chapter 4

Dyson's hierarchical model

The purpose of this chapter is only to give some background knowledge, to prepare the reader for the next chapters. It is a short introduction to the model, its properties and history, and to nucleation.

4.1 The definition

I will define *Dyson's hierarchical model* (DHM) in two cases: with a finite number N of spins and with an infinite number of spins.

4.1.1 Finite model

The spins are in both cases grouped in blocks in a hierarchical way (cf. 3.2) and the branching ratio is constant in the tree and equal to $z \geq 2$. The finite model has a finite number L of levels, so the total number of spins is $N = z^L$. The spins $s_i = \pm 1$ are numbered $i = 1, \ldots, N$ as in fig. 4.1 from left to right. The *Hamiltonian* (potential energy) is

$$H = -\frac{1}{2} \sum_{i,j=1}^{N} J z^{-\lambda d_{ij}} s_i s_j - h_0 \sum_{i=1}^{N} s_i \tag{4.1}$$

J is positive and we take it to be equal to 1. h_0 is an external uniform magnetic field. The summation convention is $\sum_{i,j=1}^{N} = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N}$. The interaction strength or coupling $J_{ij} = J z^{-\lambda d_{ij}}$ decays exponentially with ultrametric distance d_{ij} (i.e. with increasing level of the hierarchy). It is

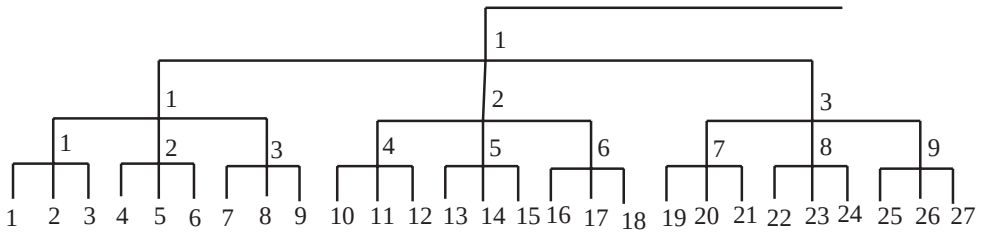


Figure 4.1: The numbering of spins and block spins for $z = 3$.

nonzero and positive (ferromagnetic) for all pairs i, j, so all spins interact, not only nearest neighbours. The positive parameter λ controls how fast the interaction decays with distance. The bigger λ the faster it falls off.

We define some *block spin* variables by adding spins in blocks: There are z^{L-l} l-level blocks and they are numbered by index $p = 1, \ldots, z^{L-l}$ from left to right as in fig. 4.1. The block spin variables are then defined by

$$S_{l,p} = \sum_j s_j, \quad (p-1)z^l + 1 \leq j \leq pz^l \tag{4.2}$$

i.e. the sum of the spins in block number p of size z^l. The definition can also be written recursively

$$S_{l,p} = \sum_{j=1}^{z} S_{l-1,z(p-1)+j} \tag{4.3}$$

On the lowest level block spins are just the spins themselves $S_{0,p} = s_p$.

All spins in a l-level block have ultrametric distance l to a spin in another l-level block, if the blocks are part of the same $l+1$-level block. So a spin interacts equally strongly with every spin in other 'neighboring' blocks.

The interaction energy part of the Hamiltonian (or the zero field Hamiltonian) H_0 can be written in two other ways that I will look at because it might give us a better understanding of the interactions. First we get the idea that neighbouring *block spins interact* on level l with strength $z^{-\lambda l}$.

$$\begin{aligned} H_0 &= -\frac{1}{2}\sum_{i,j=1}^{N} J z^{-\lambda d_{ij}} s_i s_j \\ &= -\sum_{l=0}^{L-1}\sum_{p=1}^{z^{L-l-1}}\sum_{i,j=1}^{z}\frac{1}{2} J z^{-\lambda l} S_{l,z(p-1)+i}\, S_{l,z(p-1)+j} \end{aligned} \tag{4.4}$$

For $z = 2$ this becomes

$$\begin{aligned} H_0 &= -J \sum_{l=0}^{L-1} 2^{-\lambda l} \sum_{p=1}^{2^{L-l-1}} S_{l,2(p-1)+1}\, S_{l,2(p-1)+2} \\ &= -J \sum_{l=0}^{L-1} 2^{-\lambda l} \sum_{p=1}^{2^{L-l-1}} S_{l,2p}\, S_{l,2p-1} \end{aligned} \tag{4.5}$$

which is just eq. 3.2 in [1]. This expression for $z = 2$ becomes easier to see if we write it as

$$\begin{aligned} H_0 &= -J[s_1 s_2 + s_3 s_4 + \ldots] \\ &\quad -J2^{-\lambda}[(s_1 + s_2)(s_3 + s_4) + (s_5 + s_6)(s_7 + s_8) + \ldots] \\ &\quad -J2^{-2\lambda}[(s_1 + s_2 + s_3 + s_4)(s_5 + s_6 + s_7 + s_8) + \ldots] - \ldots \end{aligned} \tag{4.6}$$

which we also find in [1, 2].

We can also use the idea that *each block acts like a mean field model*, in which all its z block spins of the level below interact equally with each other, with strenght $Jz^{-\lambda(l-1)}$. If we square the block spin to obtain 'all interact' then we need a correction for multiplying lower level block spins with themselves:

$$\begin{aligned} H_0 &= \sum_{l=1}^{L} Jz^{-\lambda(l-1)} \sum_{p=1}^{z^{L-l}} \frac{1}{2}(S_{l,p})^2 - \sum_{l=0}^{L-1} Jz^{-\lambda l} \sum_{p=1}^{z^{L-l}} \frac{1}{2}(S_{l,p})^2 \\ &= \sum_{l=0}^{L} Jz^{-\lambda l}(z^{\lambda}(1 - \delta_{0l}) - (1 - \delta_{lL})) \sum_{p=1}^{z^{L-l}} \frac{1}{2}(S_{l,p})^2 \end{aligned} \tag{4.7}$$

For $z = 2$ this will look like Dyson's original definition [Dyson, eq. 3.3].

4.1.2 Infinite model

There are at least two possible ways to define an *infinite system*. The first way is repeatedly to replace every spin with a block of z spins. The tree grows from its endpoints and branches an infinite number of times. Then there will be a single top ancestor node. Every ancestor has a finite number of ancestors but an infinite number of descendants. The second way is to add $(z-1)$ blocks of the size of the system to form a block on the next level again

and again. The tree grows from its top ancestor node an infinite number of times. Then every ancestor node has a finite number of descendants but an infinite number of ancestors. The ultrametric distance between two arbitrary spins in this case is well defined. Both ways have an infinite number of spins and levels, but the models are different. We need the model that naturally corresponds to the thermodynamic limit $L \to \infty$ of the finite models. That will be the second way! The spins are then numbered $i = 1, 2, 3, \ldots$. The block spins are defined the same way as before and numbered $p = 1, 2, 3, \ldots$. Equations for the Hamiltonian can be obtained by replacing L with ∞ in the finite model equations.

4.2 Long range forces: The history of the DHM

λ is how fast the interaction decays with UM distance. In the limit $\lambda \to \infty$ we get $J_{ij} = J$ for i and j in the same 1-level block while J_{ij} is zero for longer distances. In that limit the 1-level blocks are independent systems. The mean field limit is $\lambda = 0$ where $J_{ij} = J$ for all pairs i, j.

The spins of an Ising model of a ferromagnet should of course be embedded in a euclidian space. Dyson introduced the DHM in studying a one dimensional Ising ferromagnetic chain. In this case the spins were placed on a 1-D 'lattice', as in fig. 4.1. Originally Dyson's model had $z = 2$. It is often refered to in the literature as 'the hierarchical model' (HM). As a 1-D model DHM simulates *long range power law* interactions: The euclidian distance is $r_{ij} = |i - j|$ and $z^{d_{ij}} \leq r_{ij} \leq z^{d_{ij}} - 1$. This means $z^{-\lambda d_{ij}} \geq r_{ij}^{-\lambda} > z^{-\lambda(d_{ij}+1)}$. The interaction goes like $J_{ij} = Jz^{-\lambda d_{ij}} \approx Jr_{ij}^{-\lambda}$ i.e. a power law. The DHM thus simulates a system with a power law potential. Dyson used the DHM to prove the existence of a phase transition (spontaneous magnetization at finite temperature) for $1 < \lambda < 2$ for the power-law chain.

Another thing also makes the DHM interesting. Baker [3] found that DHM is a model where Wilson's renormalisation group procedure can be applied exactly. He analyzed the critical properties (see also [4]). He also extended the results to higher dimensions: In fig. 4.2 we see some examples of how the spins of DHM can be homogeneously distributed in higher dimensional spaces. In higher dimensions the number of spins in a block is

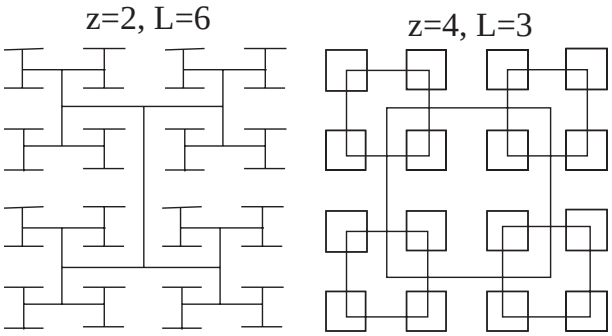


Figure 4.2: Spins of DHM's distributed homogeneously in 2D space

proportional to the sphere-like volume it occupies, i.e. $r^D \sim z^d$ where r and d are the 'sizes' of the 'sphere'. The interaction falling off exponentially with d thus falls off algebraically with r: $z^{-\lambda d} \sim r^{-\lambda D}$ (a power law).

Other relevant references are [5] ($\lambda = 1.5$), [6] (random interactions), [2] (random field) and [1] (Parisi function).

4.3 The maximal 'force'

Let us get a better understanding of the interactions. I will look at the quantity $\sum_{j \neq i} J_{ij}$ which is the maximal 'force' that can be exerted on spin i by the rest of the spins. This is also equal to the critical temperature (Weiss) and h_{max} defined in app. B. To do the sum $\sum_{j \neq i} J_{ij}$ we can sum over levels: On each level l there are $z-1$ neighboring blocks of z^l spins to the l-level block that spin i belongs to (See fig. 4.3). Spin i interacts equally with spins in these neighboring blocks. The total interaction from a neighboring block of z^l spins is therefore $z^l J z^{-\lambda l} = J z^{(1-\lambda)l}$. We see that for $\lambda < 1$ the interaction with a block grows with UM distance, so that the large distant blocks are dominating. For $\lambda > 1$ it decays with UM distance, so that small nearby blocks dominate. For $\lambda = 1$ we see that a l-level block at UM distance l counts as much as a single spin at UM distance zero. The sum is

$$\sum_{j \neq i} J_{ij} = \sum_{l=0}^{L-1} J(z-1) z^{(1-\lambda)l} = J(z-1) \begin{cases} \frac{1-z^{(1-\lambda)L}}{1-z^{1-\lambda}} & for \quad \lambda \neq 1 \\ L & for \quad \lambda = 1 \end{cases} \tag{4.8}$$

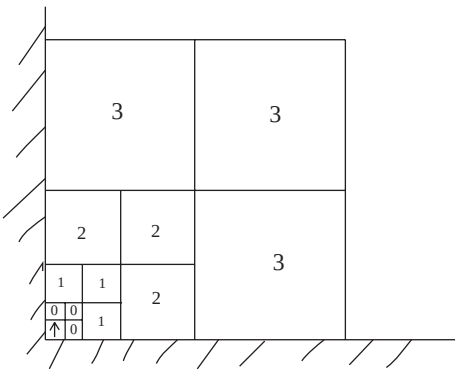


Figure 4.3: Neighbouring blocks to an arbitrary spin i (the arrow) for $z = 4$. The numbers show the levels.

For an infinite system ($L = \infty$) we have

$$\sum_{j \neq i} J_{ij} = \sum_{l=0}^{\infty} J(z-1) z^{(1-\lambda)l} \tag{4.9}$$

Define the function

$$f_{L,z}(\lambda) = \sum_{l=0}^{L-1} z^{(1-\lambda)l} = \begin{cases} \frac{1-z^{(1-\lambda)L}}{1-z^{1-\lambda}} & for \quad \lambda \neq 1 \\ L & for \quad \lambda = 1 \end{cases} \tag{4.10}$$

for $L = 1, 2, \ldots, \infty$. Then

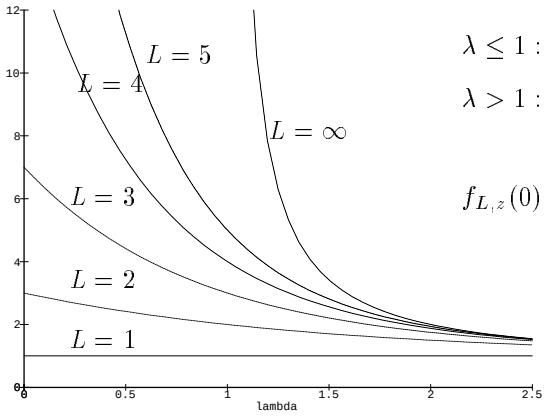
$$\sum_{j \neq i} J_{ij} = h_{max} = T_c^{Weiss} = J(z-1) f_{L,z}(\lambda) \tag{4.11}$$

This function is shown in fig. 4.4 for $L = 1$ *to* 5 and ∞. We see that T_c^{Weiss} grows when λ becomes smaller. This is because the interactions become stronger for smaller λ, so the thermal fluctuations must be stronger to compete with them.

For $\lambda \leq 1$ we have $f_{\infty,z}(\lambda) = \infty$. This means that an infinite energy increase associated with a spin flip is possible (see app. B). If we want to have a well behaved thermodynamic limit for $\lambda \leq 1$ then we can do a little trick that one does for the infinite range interactions (mean field model) at $\lambda = 0$: Here we must scale down the interactions like $J_{ij} \sim 1/N$. For DHM we therefore can set

$$J = J_0 z^{-(1-\lambda)L} = J_0 N^{-(1-\lambda)} \quad for \quad \lambda < 1 \tag{4.12}$$

$$J = J_0 L^{-1} \quad for \quad \lambda = 1 \tag{4.13}$$



$$\lambda \leq 1: \; f_{L,z}(\lambda) \to \infty \;\; for \;\; L \to \infty$$

$$\lambda > 1: \; f_{L,z}(\lambda) \to \frac{1}{1-z^{1-\lambda}} \;\; for \;\; L \to \infty$$

$$f_{L,z}(0) = \frac{z^L - 1}{z-1}, \quad f_{L,z}(1) = L, \quad f_{L,z}(2) = \frac{z - z^{1-L}}{z-1}$$

$$f_{L,z}(\lambda) \to 1 \;\; for \;\; \lambda \to \infty$$

$$f_{1,z}(\lambda) = 1$$

Figure 4.4: The function $f_{L,z}(\lambda)$ for $z = 2$.

for the finite models, so that h_{max} does not diverge:

$$\begin{aligned} h_{max} &= J(z-1)\frac{1 - z^{(1-\lambda)L}}{1 - z^{(1-\lambda)}} \\ &= J_0(z-1)\frac{z^{-(1-\lambda)L} - 1}{1 - z^{(1-\lambda)}} \\ &\to J_0(z-1)\frac{-1}{1 - z^{(1-\lambda)}} < \infty \quad for \;\; L \to \infty \end{aligned} \tag{4.14}$$

What we do is only to change our glasses to see what goes on when $L \to \infty$.

4.4 Stability of blocks

We saw in simulations that a block uniformly filled with spins in one direction, let us say up, can be stable. The up spins in the block support each other against outside influence. It must be a condition that λ is big enough so that interactions over short distances are strong enough.

A block of z^n up spins in an infinite system will be stable for all configurations at $h_0 = 0$ (and $T = 0$) if it is stable in a uniform background of down spins. Let s_i be a spin in the block. Then the condition is

$$0 < h_i = \sum_{l=0}^{n-1} J(z-1)z^{(1-\lambda)l} - \sum_{l=n}^{\infty} J(z-1)z^{(1-\lambda)l}$$

$$
\begin{aligned}
&= \left(\sum_{l=0}^{n-1} - \sum_{l=0}^{\infty} + \sum_{l=0}^{n-1}\right) J(z-1)z^{(1-\lambda)l} \\
&= J(z-1)\left(2\frac{1-z^{(1-\lambda)n}}{1-z^{1-\lambda}} - \frac{1}{1-z^{1-\lambda}}\right) \qquad (4.15)
\end{aligned}
$$

which reduces to $0 < 1 - 2z^{(1-\lambda)n}$ or

$$\lambda > 1 + \frac{1}{n}\log_z 2 \equiv \lambda_n^* \qquad (4.16)$$

so λ has to be bigger than a limit $\lambda_n^* > 1$ where the limit must be bigger the smaller the block is. $\lambda_n^* \to 1$ for $n \to \infty$.

For $\lambda < 1$ a finite block can never beat the strongest farthest blocks. Instead we can ask when a fraction of the system is stable in the thermodynamic limit. And in order to take the limit we use eq. 4.12. So let there be L levels and let a block on level $L-m$ be filled with up spins. This block is always stable at $h_0 = 0$ and $T = 0$ if

$$
\begin{aligned}
0 < h_i &= \sum_{l=0}^{L-m-1} J(z-1)z^{(1-\lambda)l} - \sum_{l=L-m}^{L-1} J(z-1)z^{(1-\lambda)l} \\
&= J(z-1)\left(2\sum_{l=0}^{L-m-1} - \sum_{l=0}^{L-1}\right) z^{(1-\lambda)l} \\
&= J(z-1)\left(2\frac{1-z^{(1-\lambda)(L-m)}}{1-z^{1-\lambda}} - \frac{1-z^{(1-\lambda)L}}{1-z^{1-\lambda}}\right) \\
&= J_0(z-1)\left(\frac{z^{-(1-\lambda)L} - 2z^{-(1-\lambda)m} + 1}{1-z^{1-\lambda}}\right) \qquad (4.17)
\end{aligned}
$$

which reduces to $0 < z^{-(1-\lambda)L} - 2z^{-(1-\lambda)m} + 1 \to -2z^{-(1-\lambda)m} + 1$ for $L \to \infty$ or

$$\lambda > 1 - \frac{1}{m}\log_z 2 \equiv \lambda_m' \qquad (4.18)$$

This is very similar to eq. 4.16. We still have that λ must be big enough and λ_m' is bigger the smaller the block is. An interesting border between the two kinds of systems is $\lambda = 1$.

4.5 Classical nucleation theory

Our goal is to investigate the process by which the system gets from the metastable phase over the barrier to the stable phase, where the magnetization and external field have the same sign. This is a *nucleation process*. I will first describe some simple classical nucleation theory (CNT) because my theory is analogous to it [9, 10].

Fluctuations in the system create small droplets that have the stable phase inside. Many droplets quickly disappear but some will grow until they cover the whole system. This growth of large droplets that will not disappear is called *domain growth*.

In CNT we assume that the free energy density inside the droplet is that of the stable phase while the free energy density outside is that of the metastable phase. We must in other words have that the system has had time to settle in the metastable equilibrium, i.e. that the field is not too big. We also assume that the density of droplets is low so that they are non-interacting, and that we are away from the critical point so that they are compact (not fractal). The free energy cost of creating a droplet of radius R is

$$\Delta F = AR^{D-1} - BR^D \tag{4.19}$$

The positive term AR^{D-1} is the cost to having a boundary between the two phases and is proportional to the surface area. A is a surface tension. The negative term $-BR^D$ is a volume term, where $B = f_{ms} - f_s$ is the difference in free energy density between the metastable and the stable phase. B is proportional to the magnitude of the external field h_0. In fig. 4.5 we see that small droplets most probably will decay because it lowers the free energy, while large droplets will tend to grow.

The critical droplet size R_c is obtained by

$$0 = \frac{d}{dR}\Delta F = A(D-1)R^{D-2} - BDR^{D-1} \tag{4.20}$$

whhich gives us

$$R_c = \frac{A(D-1)}{BD} \sim \frac{1}{|h_0|} \tag{4.21}$$

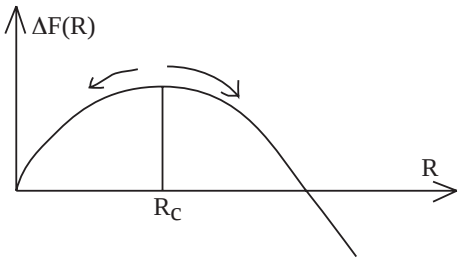


Figure 4.5: Free energy cost of droplet of radius R.

and

$$\Delta F(R_c) = \frac{1}{B^{D-1}} \left\{ A \left(\frac{A(D-1)}{D} \right)^{D-1} - \left(\frac{A(D-1)}{D} \right)^{D} \right\} \sim \frac{1}{|h_0|^{D-1}} \quad (4.22)$$

$\Delta F(R_c)$ is the *free energy barrier* that the system must overcome to get out of the metastable phase.

However we have 'forgotten' that the free energy cost to having a droplet of size R should include an entropy term associated with the various positions the droplet can have in the volume V of the system:

$$\Delta F_V(R_c) = \Delta F(R_c) - \frac{1}{\beta} \ln V \quad (4.23)$$

The mean time spent in the metastable state before a critical droplet appears is

$$\tau \propto e^{\beta \Delta F_V} = \frac{1}{V} e^{\beta \Delta F} \quad (4.24)$$

The rate of appearance of critical droplets is

$$\tau^{-1} \propto V e^{-\beta \Delta F} \quad (4.25)$$

The *nucleation rate* is the number of critical droplets per volume per unit time and is therefore proportional to $e^{-\beta \Delta F}$.

4.6 1-D long range force nucleation

Let us look at a droplet model of nucleation in the infinite Ising chain with power law potential

$$H = -\frac{1}{2}\sum_{i,j\in\mathcal{Z}} |i-j|^{-\lambda} s_i s_j - h_0 \sum_{i\in\mathcal{Z}} s_i \tag{4.26}$$

where $\mathcal{Z}$ are the integers $\ldots, -1, 0, 1, \ldots$. In section 4.2 we saw that DHM simulates a chain where $i \in \{1, 2, \ldots\}$, i.e. *positive* integers. We are going to perform a calculation that we will repeat for DHM in the next chapter. We will find the energy cost E_n of creating a compact droplet of up spins of size n in a uniform background of down spins. The external field h_0 is positive. If $E(n)$ is the energy of the configuration with a droplet of size n then the cost is $E_n = E(n) - E(0)$. If we choose the energy so that $E(0) = 0$ then we can talk about 'energies' or 'energy cost' as we like. A compact cluster is a cluster of neighbouring up spins. The size or length is the number of up spins in the droplet. First we will calculate the increase in energy when one of the two neighbouring down spins to a droplet of size n is flipped, i.e. when we go from having a droplet of size n to one of size $n+1$. Let the flipping spin have index i.

$$\begin{aligned}
\Delta E_n &= \Delta E(n \to n+1) = 2 s_i^{before} h_i^{eff} \\
&= -2\left(h_0 + \sum_{j=1}^{\infty} j^{-\lambda} s_{i+j} + \sum_{j=1}^{\infty} j^{-\lambda} s_{i-j}\right) \\
&= -2\left(h_0 - 2\sum_{r=n+1}^{\infty} r^{-\lambda}\right)
\end{aligned} \tag{4.27}$$

because the n nearest spins on both sides of s_i cancel each other. We can approximate the sum with an integral ($\lambda > 1$)

$$\sum_{r=n+1}^{\infty} r^{-\lambda} \approx \int_{n+1}^{\infty} r^{-\lambda} dr = (\lambda-1)(n+1)^{1-\lambda} \tag{4.28}$$

The energy of a configuration with a droplet of size n is

$$E_n = \sum_{m=0}^{n-1} \Delta E_m = -2h_0 n + 4(\lambda-1)\sum_{m=0}^{n-1} (m+1)^{1-\lambda} \tag{4.29}$$

Again we approximate the sum with an integral

$$\sum_{m=0}^{n-1}(m+1)^{1-\lambda} \approx \int_1^n m^{1-\lambda}dm = \frac{n^{2-\lambda}}{2-\lambda} - \frac{1}{2-\lambda} \tag{4.30}$$

and we get

$$E_n = -2h_0 n + 4\frac{\lambda-1}{2-\lambda}n^{2-\lambda} - 4\frac{\lambda-1}{2-\lambda} \tag{4.31}$$

If we regard n as a continuous variable then the energy as a function of n has qualitatively the same shape as $\Delta F(R)$ in fig. 4.5. We find the maximum by setting the derivative equal to zero or setting $\Delta E_n = 0$ and we get a critical droplet size

$$n_c = \left(\frac{h_0}{2(\lambda-1)}\right)^{\frac{1}{1-\lambda}} \sim h_0^{\frac{1}{1-\lambda}} \tag{4.32}$$

and a maximum energy

$$E_c = \left\{-2A + 4\frac{\lambda-1}{2-\lambda}A^{-\lambda+2}\right\} h_0^{\frac{2-\lambda}{1-\lambda}} + 4\frac{\lambda-1}{2-\lambda} \sim h_0^{\frac{2-\lambda}{1-\lambda}} \tag{4.33}$$

where the expression in the brackets is a constant and $A = (2\lambda-2)^{\frac{1}{1-\lambda}}$. We see that the smaller the field h_0 is, the bigger is the critical droplet. Compare with eq. 4.21 and 4.22: We see that the scaling with $h_0 \to 0$ of critical size and energy/free energy cost is the same when $D = 1$ and $\lambda = 2$ (when forces are short-ranged).

Chapter 5

The energy landscape

Free energy barriers determine nucleation rates through an exponential expression $e^{-\beta B_F}$ where B_F is the free energy barrier. At low temperatures the nucleation rate is proportional to $e^{-\beta B_E}$ where B_E is the energy barrier, since the free energy $F = E - ST$ becomes equal to the energy when $T \to 0$. The prefactor takes care of entropic effects. In this chapter we will study energy barriers and the energy landscape and leave entropic effects to the next chapter.

5.1 Droplets in Dyson's hierarchical model

We saw in the previous chapter that classical nucleation theory considers *compact droplets*. I will define a compact droplet of up spins in a uniform background of down spins in DHM. Intuitively the up spins must be packed together to occupy the smallest volume (whatever that means in an ultrametric space). A droplet can be grown by flipping up one of the down spins at a time. The following rule must be followed: Flip one of the unflipped (down) spins with the smallest UM distance to the previously flipped spin. After having grown a droplet of some size, the state of the system will be characterised by the fact, that there is at most one block on each level which is not completely filled with either up or down spins. One example of growing a droplet is simply to flip spins in the order they are numbered from left to right as in fig. 4.1. This will comply with the above definition. See a compact droplet of size 23 in fig. 5.1. A configuration of the system where the

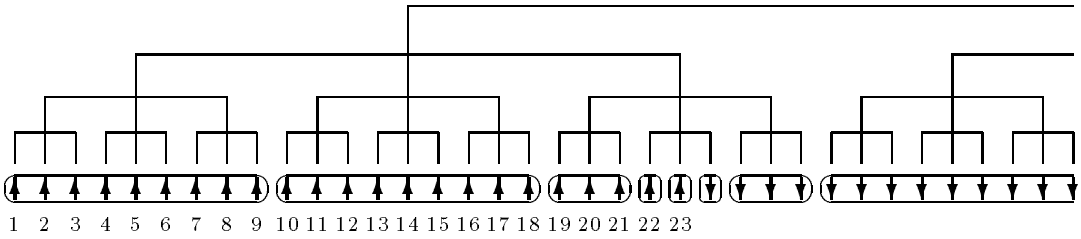


Figure 5.1: The compact droplet $n = 23 = (212)_3$ for DHM with $z = 3$. The numbering of spins is an optimal sequence of their flips. The partition into blocks is shown.

up spins constitute a compact droplet will be called an *optimal state*. The droplet growth rule does not uniquely determine the sequence for flipping spins. For example in fig. 5.1 when spin 19 was flipped we could just as well have flipped one of the spins 20-27. Therefore there may be more than one optimal state at a given magnetization (the droplet size or number of up spins n corresponds to magnetization). This degeneracy simply means that a droplet can be located at different positions in the system.

Consider the different configurations with n up spins. The optimal states have the minimum energy of all these configurations, because packing spins together in droplets minimize the energy cost of violated bonds[1]. Optimal states are thus ground states at a given magnetization. All optimal states with a given droplet size n have the same energy E_n.

Define A_- and A_+ to be the two uniform configurations of down and up spins respectively. They are the two ground states of the system when $h_0 = 0$. Growing a droplet corresponds to tracing out a single-spin-flip path in configuration space from A_- to A_+. We will call such a path an *optimal path*. There are many optimal paths because they are not uniquely determined by the rule. We will refer to the n'th state on an optimal path as the *state* n. The function E_n gives the energies of the states along an optimal path (energy profile), and is the same for all optimal paths. In the next section we will calculate E_n. Single-spin-flip paths that involve non-optimal states are called *suboptimal paths*.

[1]This is also the case with long range ferromagnetic interactions in euclidian space. In the DHM case I have not managed to give a mathematical proof of this, but I am content with my intuitive physical feeling of it.

The states of an Ising system with N spins are the corners of an N-dimensional hypercube and single-spin-flip dynamics is a diffusion on this hypercube, with the edges as possible transitions. With the energy of each state we have diffusion in an energy landscape. I define the energy barrier $B_E(\alpha, \beta)$ between two arbitrary states α and β to be the least energy required to get from α to β, or more precisely:

$$B_E(\alpha, \beta) = \min_{\sigma(\alpha,\beta)} \max_{\gamma \in \sigma(\alpha,\beta)} (E(\gamma) - E(\alpha)) \tag{5.1}$$

where $\sigma(\alpha, \beta)$ denotes a path from α to β. The energy barrier between A_- and A_+ is important for nucleation in DHM, since it can be used to predict the nucleation rate. We will study nucleation with $h_0 > 0$ so that the metastable phase has negative magnetization and the stable phase positive magnetization and A_+ is the only ground state of the system. Then $B_E(A_-, A_+)$ is the appropriate barrier. Of course we can choose the signs the other way—the two cases are symmetric. I define[2] $E(A_-) = 0$ so that $B_E(A_-, A_+) = \min_\sigma \max_{\gamma \in \sigma} E(\gamma)$.

The minimization over paths in eq. 5.1 will simply give an optimal path for σ, because if there existed a path from A_- to A_+ with a lower maximum energy, the optimal states would not all be ground states at their given magnetizations. Therefore the energy barrier $B_E(A_-, A_+)$ is simply the maximum energy along an optimal path. This barrier will be called E_{max}, i.e.

$$E_{max} = \max_n E_n \tag{5.2}$$

We can use the same argument to find the energy barrier between two arbitrary optimal states n and m on the same optimal path. It is just

$$B_E(n, m) = \max_j (E_j - E_n) \tag{5.3}$$

where j runs between n and m.

The optimal path is energetically optimal. But it is also the shortest path between its states. Furthermore, the lowest energy increase associated with flipping an up or down spin in an optimal state is realized when making a transition to a predecessor or succesor respectively on the optimal path. We call such transitions *optimal transitions*.

[2] We can choose the zero point of the energy as we wish.

With the knowledge of optimal paths and the function E_n it is easiest to think of the nucleation problem as a diffusion process that starts in A_- at time $t = 0$. The problem studied in CNT is slightly different because we start in the metastable phase at time $t = 0$. But in both cases it is the same nucleation physics that takes the system over the barrier, and the barrier height is basically the same. So the problem I will solve is diffusion starting in A_- at time $t = 0$.

5.2 Energy calculation

The DHM has a symmetry similar to translational invariance in lattice models: Two subtrees of the same ancestor node can be swapped without changing the metric and energy. Given an optimal path it is possible to transform it by swapping subtrees into the optimal path corresponding to flipping spins from left to right[3]. This is why all optimal paths have the same energy profile E_n. And therefore we can simply think of a picture like fig. 5.1 when thinking about an optimal path.

In this section we will calculate the function E_n for both finite and infinite L, and since $E(A_-) = 0$ it is

$$E_n = \sum_{s=0}^{n-1} \Delta E(s \to s+1) \tag{5.4}$$

where $\Delta E(s \to s+1) \equiv \Delta E_s$ is the energy increase in going from state s to state $s+1$, i.e. in flipping spin number $s+1$. According to eq. B.3 it is

$$\Delta E_s = -2h_{s+1}^{eff} = -2h_0 - 2 \sum_{j \neq s+1} s_j J_{j,s+1} \tag{5.5}$$

because the spin is -1 before it is flipped.

A number n can be written in the number system with base z:

$$n = (\ldots c_l^n c_{l-1}^n \ldots c_0^n)_z = \sum_{l=0}^{\infty} c_l^n z^l \tag{5.6}$$

[3]As a consequence, in an optimal state in a finite system not only the up spins will be gathered in a compact droplet but also all the down spins, see fig. 5.1.

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
c_0^n	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0
c_1^n	0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3	0	0	0	0	1
c_2^n	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
c_3^n	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.1: Digits or numbers of flipped blocks for $z = 4$

where the digits c_l^n are given by[4]

$$c_l^n = (n \text{ div } z^l) \bmod z \tag{5.7}$$

and $0 \leq c_l^n \leq z-1$, see table 5.1. If $n \leq z^L$ then higher digits than c_L^n will all be zero and we can skip them:

$$n = (c_L^n c_{L-1}^n \dots c_0^n)_z = \sum_{l=0}^{L} c_l^n z^l \tag{5.8}$$

The systematic way of counting, using digits c_l^n with number base z as seen in table 5.1, represents the filling of blocks as the droplet grows: Consider state n and spin $n+1$ which is the next to be flipped. As I remarked in the beginning of this chapter, there is at most one block on each level which is not completely filled with either up or down spins. If there is such a block on level l then spin $n+1$ must belong to it. The $z-1$ neighboring blocks of z^l spins that spin $n+1$ does not belong to (see fig. 4.3), are therefore all uniformly filled. The number of these blocks that are filled with up (flipped) spins in state n is simply the digit c_l^n. Then the number of these blocks filled with down spins is $z-1-c_l^n$. The digits $(c_L^n, \dots, c_0^n)_z$ prescribe a *partition* of the droplet and of the whole system into uniform blocks chosen as big as possible. This partition is shown in fig. 5.1 for $n = 23 = (212)_3$.

It turns out that the sum $\sum_{j \neq s+1} s_j J_{j,s+1}$ that we need in eq. 5.5 can be expressed in terms of the digits c_l^s for state s. There are c_l^s neighboring blocks to spin $s+1$ of z^l up spins, and $z-1-c_l^n$ neighboring blocks of z^l down spins. The magnitude of the interaction from such a neighboring l-level block is $Jz^{(1-\lambda)l}$ (see section 4.3). Our considerations applies so far to both finite and infinite systems. According to our considerations the sum is for a finite system with L levels

[4]With a and b integers, a div b denotes division with truncation and $a \bmod b$ is the remainder of division.

$$\begin{aligned}\sum_{j\neq s+1} s_j J_{j,s+1} &= \sum_{l=0}^{L-1} c_l^s J z^{(1-\lambda)l} - \sum_{l=0}^{L-1}(z-1-c_l^s) J z^{(1-\lambda)l} \\ &= -J\sum_{l=0}^{L-1}(z-1-2c_l^s) z^{(1-\lambda)l} \qquad (5.9)\end{aligned}$$

For an infinite system this equation also holds if we set $L=\infty$ and have infinite sums instead. So for $L=1,2,\ldots,\infty$ we get

$$\begin{aligned}\Delta E_s &= -2h_0 + 2J\sum_{l=0}^{L-1}(z-1-2c_l^s) z^{(1-\lambda)l} \\ &= -2h_0 + 2J(z-1) f_{L,z}(\lambda) - 4J\sum_{l=0}^{L-1} c_l^s z^{(1-\lambda)l} \\ &= -2h_0 + 2h_{max} - 4J\sum_{l=0}^{L-1} c_l^s z^{(1-\lambda)l} \qquad (5.10)\end{aligned}$$

where eq. 4.10 was used. In fig. 5.2 ΔE_s is plotted as function of s for $h_0 = 0$.

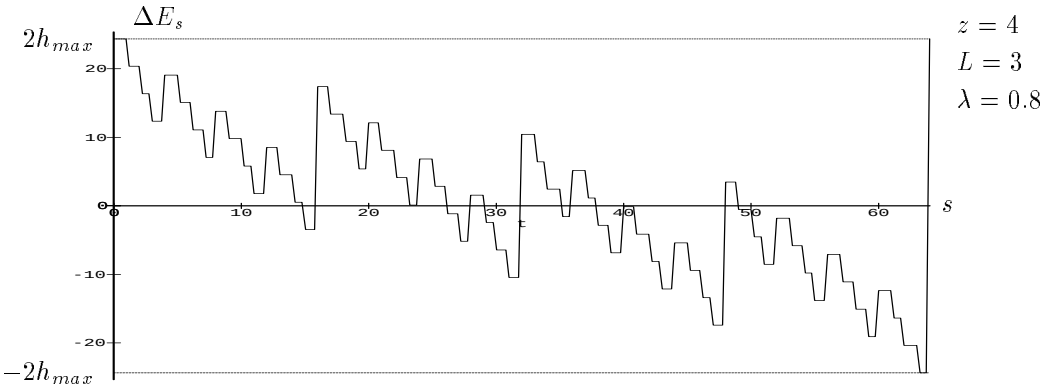


Figure 5.2: ΔE_s for the spin flips at $h_0 = 0$. The two bounds $\pm 2h_{max}$ are shown.

In agreement with App. B.1 we see that the spin flips with highest and lowest energy increases are the ones that flip a spin in a uniform background of spins (states $s=0$ and $s=N$). And we see that at $h_0 = 0$ we have $|\Delta E_s| \leq 2h_{max}$.

5.2.1 Summation

With eq. 5.10 we can now perform the summation eq. 5.4

$$E_n = \sum_{s=0}^{n-1} \Delta E_s = -2nh_0 + 2J \sum_{l=0}^{L-1} \left\{ z^{(1-\lambda)l} \sum_{s=0}^{n-1} (z - 1 - 2c_l^s) \right\} \tag{5.11}$$

Let us first do the sum over s on the right hand. The c_l^s cycles through the values $0, \ldots, z-1$ as s increases (see table 5.1). It increases by 1 whenever s is divisible by z^l, except when s is divisible by z^{l+1} where it jumps back to 0. It takes on each value z^l times before it makes a step. Accordingly the term $(z - 1 - 2c_l^s)$ cycles from $+(z-1)$ to $-(z-1)$ in steps of –2, making a step whenever s is divisible by z^l and jumping back to $z-1$ whenever s is divisible by z^{l+1}. We split the sum into three pieces:

$$\sum_{s=0}^{n-1} = \sum_{s=0}^{r-1} + \sum_{s=r}^{u-1} + \sum_{s=u}^{n-1}$$

where $r = n - (n \bmod z^{l+1})$ and $u = n - (n \bmod z^l)$ and $r \leq u \leq t$. The terms in the first sum cancel because r is divisible by z^{l+1} so that the terms go from $z-1$ to $-(z-1)$ a whole number of times . In the second sum c_l^s will go from 0 to $c_l^n - 1$ and it will take on each value z^l times because u is divisible by z^l. So this sum is

$$\begin{aligned} \sum_{s=r}^{u-1} z - 1 - 2c_l^s &= z^l \sum_{j=0}^{c_l^n - 1} z - 1 - 2j \\ &= z^l (c_l^n (z-1) - 2\frac{c_l^n (c_l^n - 1)}{2}) \\ &= z^l c_l^n (z - c_l^n) \end{aligned}$$

The last sum has a constant summand $= (z - 1 - 2c_l^n)$ so it is

$$\sum_{s=u}^{n-1} z - 1 - 2c_l^n = (z - 1 - 2c_l^n)(n \bmod z^l)$$

Putting it together in eq. 5.11 leads to

$$E_n = -2nh_0 + 2J \sum_{l=0}^{L-1} z^{(1-\lambda)l} \left\{ z^l c_l^n (z - c_l^n) + (z - 1 - 2c_l^n)(n \bmod z^l) \right\} \tag{5.12}$$

Although this expression is suitable for a *computer*, it is not easy to interpret.

I define some '*distant fields*' h_m^n, $m = 1, 2, 3, \ldots$:

$$h_m^n = h_0 - J \sum_{l=m}^{L-1} (z - 1 - 2c_l^n) z^{(1-\lambda)l} \tag{5.13}$$

This is the field on the flipping spin (number $n+1$) from all spins that are more than an UM-distance of m away, and from the external field h_0.

With these fields the E_n can be expressed in another way. Part of the right hand side in eq. 5.12 can be rewritten using $(n \bmod z^l) = \sum_{j=0}^{l-1} c_j^n z^j$:

$$\sum_{l=0}^{L-1} z^{(1-\lambda)l} (z - 1 - 2c_l^n) \sum_{j=0}^{l-1} c_j^n z^j = \sum_{l=0}^{L-1} \sum_{j=0}^{l-1} c_j^n z^j z^{(1-\lambda)l} (z - 1 - 2c_l^n)$$

Using $\sum_{l=0}^{L-1} \sum_{j=0}^{l-1} = \sum_{j=0}^{L-1} \sum_{l=j+1}^{L-1}$ (summation of triangle matrix components in different orders) and interchanging the summation-indices, this can be expessed in terms of the distant fields:

$$\begin{aligned} \sum_{j=0}^{L-1} \sum_{l=j+1}^{L-1} c_j^n z^j z^{(1-\lambda)l} (z - 1 - 2c_l^n) &= \sum_{l=0}^{L-1} c_l^n z^l \sum_{j=l+1}^{L-1} z^{(1-\lambda)j} (z - 1 - 2c_j^n) \\ &= \sum_{l=0}^{L-1} c_l^n z^l (h_0 - h_{l+1}^n) J^{-1} \end{aligned}$$

So we obtain

$$\begin{aligned} E_n &= -2h_0 \sum_{l=0}^{L-1} c_l^n z^l + 2J \sum_{l=0}^{L-1} z^{(2-\lambda)l} c_l^n (z - c_l^n) + 2 \sum_{l=0}^{L-1} c_l^n z^l (h_0 - h_{l+1}^n) \\ &= \sum_{l=0}^{L-1} 2\{J c_l^n (z - c_l^n) z^{(2-\lambda)l} - h_{l+1}^n c_l^n z^l\} \end{aligned} \tag{5.14}$$

This expression is simple enough that we can understand it geometrically together with the plots of this function. We will look at that in the next section. All results and definitions so far also hold for $L = \infty$. The terms of the sum in eq. 5.14 are zero from a certain l.

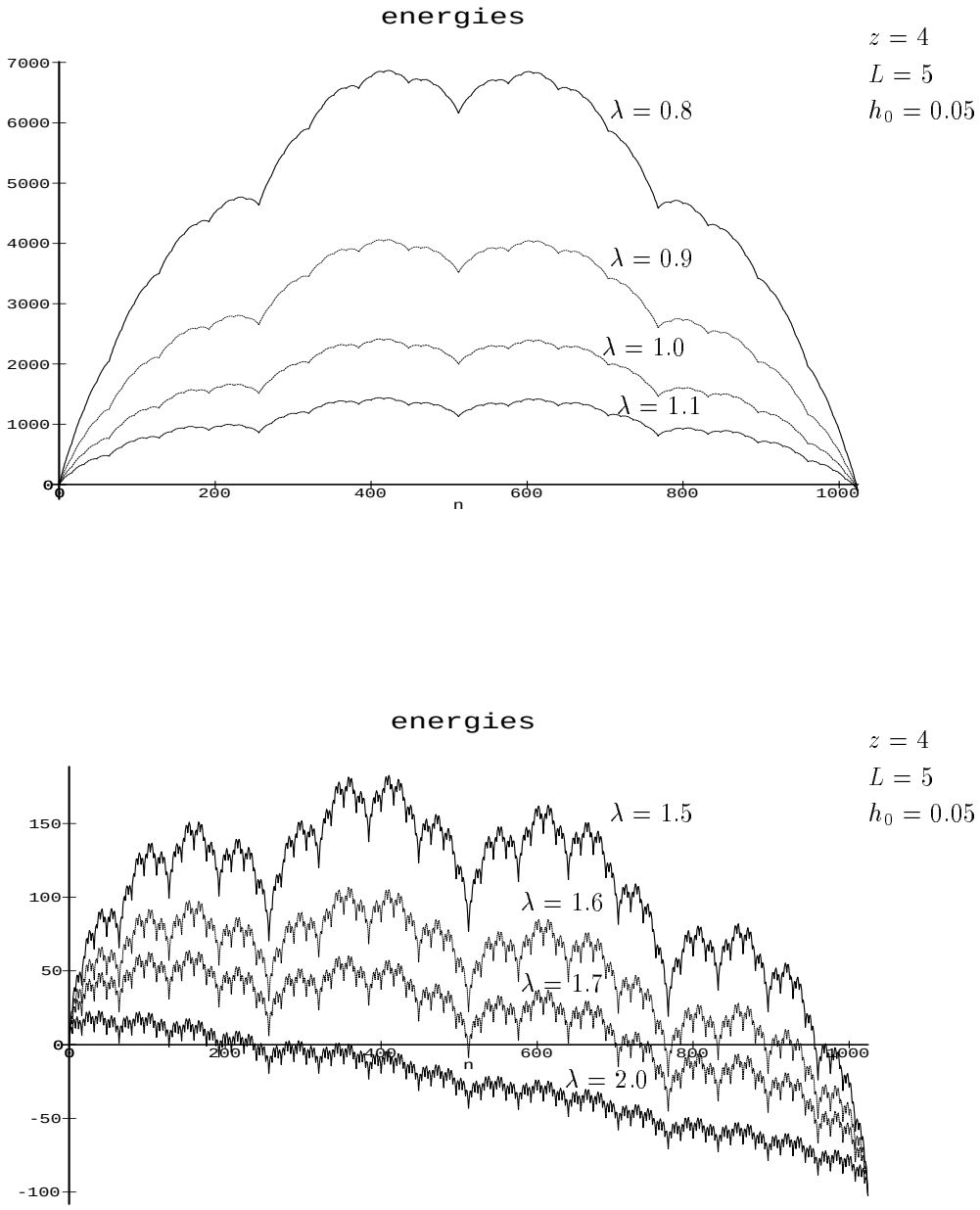


Figure 5.3: E_n for a range of λ values. All other parameters are fixed.

5.3 Energy curves

The function E_n is plotted for $z = 4,\ L = 5$ in fig. 5.3 for a range of λ-values.

For $L = \infty$ the curves have exactly the same appearance, they just go on for larger n indefinitely. For $\lambda > 1$ (and $L = \infty$) the energy E_n is finite for all n, but it may not have an upper or lower bound, it may come arbitrarily close to $\pm\infty$.

These curves should be understood in the same way as the schematic fig. 4.5 as explaining the nucleation physics, i.e. they exhibit a *barrier* E_{max}. We notice three things: First, the barrier E_{max} becomes smaller for bigger λ; second, the curves have many local minima and there are more minima for bigger λ; and third, the curves have a self-similar structure.

5.3.1 The barrier

The barrier becomes smaller for bigger λ, i.e. for shorter ranged forces. This is because the bonds that are violated by the critical droplet becomes weaker.

5.3.2 Metastable states

A *metastable state* is a configuration for which every spin flip increases the energy. A local minimum of the curves means that optimal transitions increase the energy. But optimal transitions have the lowest energy increases so a local minimum corresponds to a metastable state. From section 4.4 we know that uniform blocks corresponding to the states $n = z^k,\ k = 1, 2, \ldots$ can be stable in an infinite system. Now we see that such states can be metastable, i.e. the whole system stable. This reminds us of the energetics of filled shells/subshells in atomic/nuclear structure.[5]

There are more metastable optimal states (the curves are more rugged) for bigger λ, see for example fig. 5.4 where the curves for different λ are scaled

[5] I asked in section 3.4 whether clustering as a way to help nucleation can be compared to a *step ladder*. Now that we know the energetics we can answer these questions. The parameter λ is a measure of clustering. Clustering can be said to provide a 'ladder', namely the optimal paths which takes less energy to pass than suboptimal paths. And E_{max} becomes smaller the more clustering λ there is. But the metastable states and valleys along the optimal path before we get to the top do not work as steps/stairs. They do not influence the nucleation time scale which only depend on E_{max}.

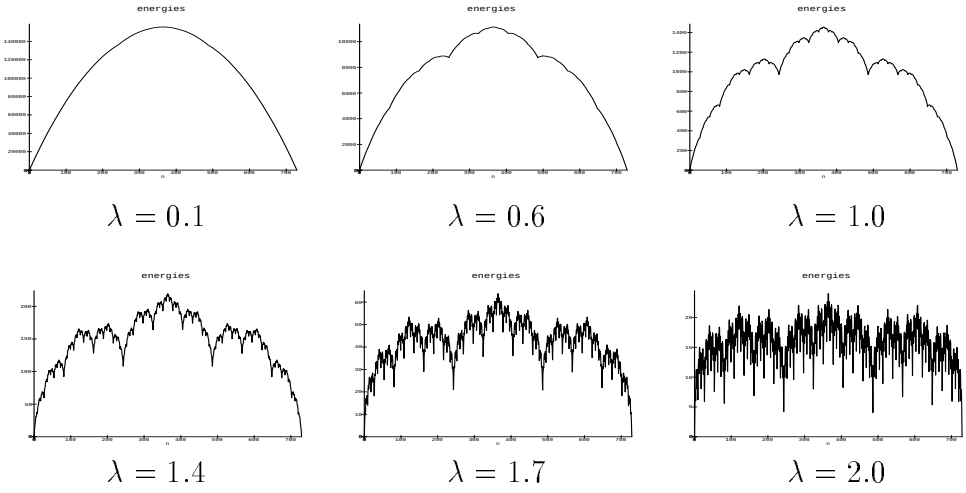


Figure 5.4: Energies for $z = 3$, $L = 6$ and zero field.

to the same size. This is what could be expected from our result in section 4.4 that stability requires a big enough λ (short enough ranged forces).

The metastable optimal states are not the only metastable states of the system, there are also suboptimal metastable states. A small calculation shows that a metastable state must consist of uniform 1-level blocks: If a non-uniform 1-level block with n_u up spins and n_d down spins is stable, then its up spins feel a positive effective field,

$$J(n_u - 1 - n_d) + y \geq 0,$$

and its down spins feel a negative effective field

$$J(n_u - n_d + 1) + y \leq 0,$$

But this is impossible to have: $J \leq J(n_u - n_d) + y \leq -J$. So in a metastable state 1-level blocks are uniform. From eq. 4.16 we then have a complete characterization of the metastable states for $\lambda > 1 + \log_z 2$ in an infinite system, namely that they are the configurations that have uniform 1-level blocks.

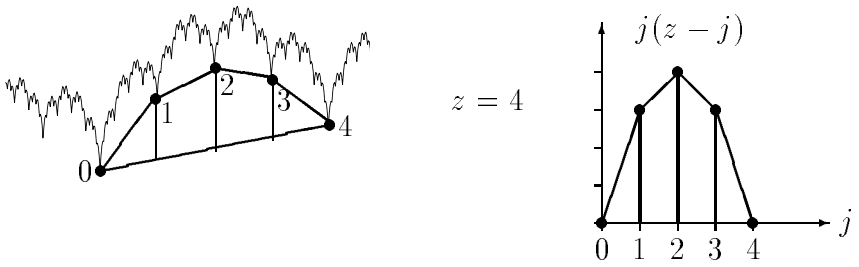


Figure 5.5: The 'parabolic' hill

5.3.3 Self-similar geometry

The graphs of E_n have a very simple self-similar structure. They are made up of parabolic "hills", each in turn made up of z times as many parabolic hills on a smaller scale etc. etc. There is a factor of $z^{2-\lambda}$ difference in heights of the hills between succesive levels. This can be seen from eq. 5.14: Choose a level m. In fig. 5.5 we examine the energy of the states $n_j = uz^{m+1} + jz^m$ where u is an arbtrary integer and $j = 0, 1, 2, 3, 4$, $(z = 4)$. We need to know the digits $c_l^{n_j}$ for these states in order to use eq. 5.14. For $l < m$ these digits are zero. For $j < 4$ we have $n_j = (\ldots x^1 x^0 j 0 \ldots 0)_z$ where $(\ldots x^1 x^0)_z = u/z^{m+1}$. From eq. 5.14 we get

$$E_{n_j} = E_{n_0} - 2h_{m+1}^{n_0} z^m j + 2j(z-j)z^{(2-\lambda)m} J \tag{5.15}$$

For $j = 4$ we can not have $c_m^{n_j} = j$. Instead we have to carry 1 to the next digits. More than one digit may be affected this way, but eq. 5.15 still holds. This may be seen in the following way: $h_{m+1}^{n_0}$ is the distant field from spins outside the block of size z^{m+1} that is being filled for $n_0 \leq n \leq n_4$. During this filling h_{m+1}^n is therefore unchanged and equal to $h_{m+1}^{n_0}$. The change in energy between n_0 and n_4 is only determined by $h_{m+1}^{n_0}$, because energy in internal interactions between the spins in the block is conserved when they are flipped simultaneously. Only interactions with the outside give a change, so

$$E_{n_4} - E_{n_0} = -2h_{m+1}^{n_0} z^{m+1} \tag{5.16}$$

with a minus because the spins are -1 before they are flipped. This shows that eq. 5.15 also holds for $j = 4$. From eq. 5.16 it is seen that $-2h_{m+1}^{n_0}$ is the slope of the line in fig. 5.5 connecting points 0 and 4. The term $j(z-j)z^{(2-\lambda)m}J$ gives a 'parabola' (see fig. 5.5) which is added to the linear term $h_{m+1}^{n_0} z^m j$. This is the hill, and the factor $z^{(2-\lambda)m}$ gives its scaling.

For $\lambda < 2$ we see decreasing heights of hills of decreasing level, while for $\lambda > 2$ the heights increase. For $\lambda \gg 2$ (which means independent 1-level blocks) the curve is a row of 1-level hills.

5.3.4 Fractals

We get a *fractal* curve if we plot energies of optimal paths in infinite systems as a function of magnetization $m \in [-1, 1]$ instead of n. It could be interesting to calculate fractal dimensions. Do they depend on λ? An optimal path is only part of the full energy landscape. What does the fractal curve tell about the energy landscape? It is an interesting mathematical task to define fractal dimensions for an energy landscape, where states are sites on a hypercube and we use the Hamming distance between them. I think our energy landscape is fractal. Fractals, scaling etc. are universal features of many complex systems, but in DHM there is nothing mysterious about it. It comes from the self-similar nature of the interactions that we defined to begin with. The fractal dimensions may be of relevance to dynamics (anomalous diffusion, ultradiffusion), but I have not pursued this further.

5.4 The energy-barrier and critical droplet

In this section I discuss the quantities

$$E_{max} = \max E_n \quad n_{max} = \arg\max E_n \tag{5.17}$$

The droplet in the state that has the maximum energy is called the *critical droplet*, and $n_{max} = (\dots c_j^{n_{max}} \dots c_0^{n_{max}})_z$ gives its composition. The two quantities depend on z, L, λ, h_0, J. I will discuss some of their properties as functions of L, h_0 and λ which are the interesting parameters in this respect.

E_{max} and n_{max} are always finite for $L < \infty$, but may not be finite when $L = \infty$. E_{max} is an increasing function of L and a decreasing function of λ (other parameters held fixed), because for any n: E_n increase if L is increased, because more terms are added in the sum eq. 5.14, and all terms increase if λ is increased. n_{max} is also increasing with increasing L. The critical droplet has to be bigger if there are more spins in the system that exerts an opposite field in it.

Both E_{max} and n_{max} are decreasing as functions of h_0: The possible n_{max} values are the n-values where the *convex hull* of the E_n-curve has a 'corner'. These corners are independent of h_0. When λ increases there will be fewer possible n_{max}-values (corners) because the curve gets more rugged. As h_0 increases one of the corners will be the maximum until at a certain point the maximum jumps to the next corner, with a lower n-value. Finally n_{max} jumps to zero at $h_0 = h_{max}$ where it stays for $h_0 > h_{max}$. The integer valued function $n_{max}(h_0)$ of h_0 is a piecewise constant, noncontinuous and decreasing function (descending stairs).

If we know the function $n_{max}(h_0)$ then we also know $E_{max}(h_0)$. Actually

$$E_{max}(h_0) = E_{n_{max}(h_0)} = -2n_{max}(h_0)h_0 + f(n_{max}(h_0)) \tag{5.18}$$

where f is some function (see eq. 5.12). $n_{max}(h_0)$ is piecewise constant so it is easy to differentiate $E_{max}(h_0)$:

$$\frac{\partial E_{max}(h_0)}{\partial h_0} = -2n_{max}(h_0) \tag{5.19}$$

$E_{max}(h_0)$ is simply the integral of $-2n_{max}(h_0)$. There is no barrier for $h_0 \geq h_{max}$ (see eq. 4.11) so $E_{max}(h_{max}) = 0$. Thus

$$E_{max}(h_0) = \int_{h_{max}}^{h_0} -2n_{max}(h)dh = 2\int_{h_0}^{h_{max}} n_{max}(h)dh \tag{5.20}$$

$E_{max}(h_0)$ is therefore a piecewise linear, continuous and decreasing function.

At $h_0 = 0$ and for $L < \infty$ it is seen on the graphs of E_n with odd z that the maximum is in the middle of the interval. More precisely $n_{max} = (d \ldots d)_z = \frac{1}{2}(z^L - 1)$ where $d = (z-1)/2$. Then $h_{l+1}^{n_{max}} = 0$ for all l and from eq. 5.14

$$E_{max} = 2Jd(z-d)\sum_{l=0}^{L-1} z^{(2-\lambda)l} = \frac{J}{2}(z^2-1)\begin{cases} \frac{1-z^{(2-\lambda)L}}{1-z^{2-\lambda}} & for \quad \lambda \neq 2 \\ L & for \quad \lambda = 2 \end{cases} \tag{5.21}$$

When z is even there is a local *minimum* at $n = \frac{1}{2}(z^L - 1)$ and there are two maxima to the left and right of $t = \frac{1}{2}(z^L - 1)$. But still

$$E_{max} \sim \sum_{l=0}^{L-1} z^{(2-\lambda)l} \tag{5.22}$$

is a good approximation. In both cases then, for fixed λ

$$E_{max} \sim z^{(2-\lambda)L} = N^{2-\lambda} \quad for \quad \lambda \neq 2 \tag{5.23}$$

and

$$E_{max} \sim \log N \quad for \quad \lambda = 2 \tag{5.24}$$

E_{max} diverges for $\lambda \leq 2$ but converges for $\lambda > 2$ when $L \to \infty$.

Let us look at $L = \infty$. For $\lambda > 2$ the barrier E_{max} is finite at $h_0 = 0$ and since it decreases with h_0 it is also finite for all $h_0 \geq 0$. For $1 < \lambda \leq 2$, E_{max} is infinite at $h_0 = 0$, but finite at $h_0 > 0$, as for $\lambda > 2$, and we will look into this now:

What is the energy of states $n' = z^l$ where the droplet is one filled l-level block? These states are $n' = (\ldots 010 \ldots 0)_z$ with $c_l^{n'} = 1$. From eq. 5.14 we get

$$\begin{aligned} E_{n'} &= 2J(z-1)z^{(2-\lambda)l} - 2h_{l+1}^{n'} z^l \\ &= 2J(z-1)z^{(2-\lambda)l} - 2z^l \left(h_0 - J \sum_{m=l+1}^{\infty} (z-1) z^{(1-\lambda)m} \right) \\ &= 2J(z-1)\frac{z^{(2-\lambda)l}}{1-z^{1-\lambda}} - 2h_0 z^l \equiv f(l) \end{aligned} \tag{5.25}$$

where $n' = z^l$. We see that when $n' \to \infty$ the negative second term will dominate the first positive term so that $E_{n'} \to -\infty$ for $n' \to \infty$. If $E_{n'} < 0$ for some n' then E_{max} must be before that, $n_{max} < n'$. The reason is that the E_n curve for $n > n'$ will have the same increases given by eq. 5.10, as it had for $n < n'$, only with some negative term added, so the E_n-curve will never be able to achieve as high values as it could for $n < n'$. Therefore n_{max} is finite and then also E_{max} finite for $h_0 > 0$ and $L = \infty$.

What happens as $h_0 \to 0$? If we regard $f(l)$ defined in eq. 5.25 as a function of a continuous variable l, then we have a continuous approximation f to the energy E_n (a lower bound: $f(\log_z n) \leq E_n$ for all n). By differentiating after l and setting equal to zero (cf. section 4.6) we find

$$\arg\max_l f = \frac{1}{\lambda - 1} \log_z \frac{(2-\lambda)(z-1)}{(1-z^{1-\lambda})h_0} \tag{5.26}$$

i.e.

$$n'_{max} = \left(\frac{(2-\lambda)(z-1)}{(1-z^{1-\lambda})h_0} \right)^{\frac{1}{\lambda-1}} \tag{5.27}$$

which is a rough approximation to the true n_{max}. And we get an approximation to E_{max}:

$$E'_{max} = 2J\frac{z-1}{1-z^{1-\lambda}}\left(\frac{(2-\lambda)(z-1)}{(1-z^{1-\lambda})h_0}\right)^{\frac{2-\lambda}{\lambda-1}} - 2h_0\left(\frac{(2-\lambda)(z-1)}{(1-z^{1-\lambda})h_0}\right)^{\frac{1}{\lambda-1}} \quad (5.28)$$

Because $1 < \lambda \leq 2$, we have that the first term will diverge and the second term go to zero when $h_0 \to 0$. So we find that the maximum energy and the critical droplet go to infinity as

$$n'_{max} \sim h_0^{\frac{1}{1-\lambda}} \quad (5.29)$$

and

$$E'_{max} \sim h_0^{\frac{2-\lambda}{1-\lambda}} \quad (5.30)$$

when $h_0 \to 0$. This is the same scaling as the one we found in eq. 4.32 and 4.33 for the Ising chain with power law potential, which Dyson studied with the DHM.

Chapter 6

The free energy landscape

In this and the following chapter we only consider finite systems.

6.1 Relaxation

We want to study the diffusion of the system in its energy landscape with A_- as the starting point at time $t = 0$. We saw in the previous chapter that we have a rugged energy landscape with many metastable states. According to the standard picture of relaxation and broken ergodicity in such systems [7], the system will explore bigger and bigger parts of the free energy landscape in the following way: On a given timescale τ the system will be confined to a *component* which is the part of state space accessible from the starting point without having to surmount a free energy barrier higher than F, where $\tau \sim e^{F/T}$ and T is the temperature. Being confined means that there is a small probability of escaping the component on that timescale. On this timescale the system is ergodic within the component, i.e. it revisits all parts of the component with frequencies approximately given by the Gibbs distribution restricted to the component. On smaller timescales the system will only have equilibrated within smaller components. We imagine components as lakes in the free energy landscape, with the water level equal to the smallest free energy barrier to escape. As time increases the water level increases logarithmically and lakes grow into bigger lakes. Disjoint lakes in different parts of the landscape merge in a hierarchical way with increasing water level [8].

This picture is being to explain dynamics of disordered systems e.g. ageing in spin glasses, but the intuition is based on diffusion in a one-dimensional state space. For some systems a 'pond-and-outlet' picture [7] applies instead, where ponds visited earlier are not revisited. But we think the standard picture applies in our case.

We do not know our free energy landscape. At low temperatures it will be close to the energy landscape for which we know some relevant barriers $B_E(n,m)$ given by eq. 5.3. The energy barrier $B_E(A_-,n)$ between A_- and n is relevant for predicting the time before the system 'explores' state n.

$$B_E(0,n) = \max_{0\leq j\leq n} E_j \tag{6.1}$$

But entropic effects play a role too since we only have dynamics at non-zero temperature. At low temperature we should at least take into account the degeneracy of optimal paths. I explain the physics of our process and the Arrhenius law in a heuristic way:

All the probability mass is in A_- at $t = 0$. Then it starts spreading out in configuration space according to the solution eq. A.9 to the master equation. It flows from states of magnetization n to states of magnetization $n+1$ (or from the surface of the radius-n ball in configuration space with center A_- to the radius $n+1$ ball). Most of this flow will go through optimal transitions: The flow from a state i to a state j is $p_i W_{ij}$ (see app. A), so the total flow through a state i is proportional to the probability mass in that state. The probability $p_i(t)$ of state i will tend to the Gibbs probability. For my considerations I can assume $p_i(t)$ to be proportional to the Gibbs probability. Most of the flow from the n-surface to the $n+1$-surface will therefore go through the optimal states n because the suboptimal states have exponentially smaller probabilities. Define A_n to be *the number of optimal states with n flipped spins.* The flow from an optimal state n is proportional to $e^{-\beta E_n}$ and the total flow from the n-surface to the next is therefore proportional to $A_n e^{-\beta E_n}$. The flow from A_- to n must be limited by the 'weakest' surface between 0 and n (a *bottleneck*) and is therefore proportional to $\min_{0\leq j\leq n} A_j e^{-\beta E_j}$. This bottleneck will be the free energy barrier. The time τ_n before the system explores states n is the time before enough probability mass has reached n and is the inverse of the flow to n:

$$\tau_n = \left(\min_{0\leq j\leq n} A_j e^{-\beta E_j}\right)^{-1} = \max_{0\leq j\leq n} A_j^{-1} e^{\beta E_j} \tag{6.2}$$

What we need is therefore to calculate A_n and we do that in the next section.

6.2 Counting optimal states

Now we calculate the number A_n of ways to place a droplet of n flipped spins, using combinatorics. Consider a state $n = (c_L^n c_{L-1}^n, \ldots, c_1^n c_0^n)_z$ with n flipped (up) spins. Consider the c_j^n j-level blocks that are filled with up spins. Assume that we have already chosen which $j+2$-level block they belong to. Now we choose which $j+1$-level block they belong to. There are $z - c_{j+1}^n$ 'empty' $j+1$-level blocks to choose among. Then we choose how to distribute the c_j^n filled j-level blocks among the z possible within the chosen $j+1$-level block. There are $\left(\begin{array}{c} z \\ c_j^n \end{array} \right)$ possible ways to do that. So the number of choices for the j-level blocks is

$$(z - c_{j+1}^n) \left(\begin{array}{c} z \\ c_j^n \end{array} \right)$$

Then we go to the level one step below and do the same arguments. The assumption this time that we already have chosen the $j+1$-level block is true since we chose that in the previous step. Do we simply get a factor $(z - c_{j+1}^n) \left(\begin{array}{c} z \\ c_j^n \end{array} \right)$ for levels $L-1$ down to 0? No, when we start at level[1] $L-1$ there is only one $j+1 = L$-level block to choose and not $z - c_{j+1}^n = z$. And when we have distributed the smallest filled blocks on level p_n where

$$p_n = \min\{i | c_i^n \neq 0\} \tag{6.3}$$

then there are no more filled blocks to distribute and no more choices to make. Multiplying the choices therefore gives

$$A_n = \prod_{j=p_n}^{L-2} (z - c_{j+1}^n) \prod_{j=p_n}^{L-1} \left(\begin{array}{c} z \\ c_j^n \end{array} \right) \tag{6.4}$$

for $n < N = z^L$. Alternatively, we can write $j = 0$ in the second product since $\left(\begin{array}{c} z \\ c_j^n \end{array} \right) = 1$ for $j < p_n$. For $n = N$ and for $n = 0$ we have $A_n = 1$ (the states A_+ and A_-). The equation also holds for $n = 0$ where $p_n = \infty$.

[1] Assume $n < z^L$. For $n = z^L$ there is only one state namely A_+ so $A_n = 1$.

A perhaps simpler way to derive A_n is to go from lower to higher levels: First we find the $p_n + 1$-level block to which the $c^n_{p_n}$ p_n-level blocks belong. There are z^{L-p_n-1} possible blocks in an empty system. Then we place the p_n-level blocks among z possible in $\left(\begin{array}{c} z \\ c^n_{p_n} \end{array} \right)$ ways. Then we go one level up. The $c^n_{p_n+1}$ filled blocks on this level now has $z-1$ empty blocks to choose among, so there are $\left(\begin{array}{c} z-1 \\ c^n_{p_n+1} \end{array} \right)$ possibilities. This holds up to level $L-1$. So for $0 < n < N$ we get

$$A_n = z^{L-p_n-1} \left(\begin{array}{c} z \\ c^n_{p_n} \end{array} \right) \prod_{j=p_n+1}^{L-1} \left(\begin{array}{c} z-1 \\ c^n_j \end{array} \right) \tag{6.5}$$

which should be identical to eq. 6.4.

Let us now ask how many ways there are to flip an additional spin (an optimal transition) given a state n. The next spin must be within the p_n+1-level block that is not completely filled. Here there are $z - c^n_{p_n}$ empty blocks of z^{p_n} spins to choose among so the number of ways 'forward' is

$$w^n_f = (z - c^n_{p_n}) z^{p_n}$$

for $0 < n < N$. For $n = 0$ there are N spins to flip next.

Let us also ask how many ways there are to 'unflip' a spin (still an optimal transition) given a state n. Then we must flip a spin in one of the $c^n_{p_n}$ p_n-level blocks. The number of ways 'backwards' is

$$w^n_b = c^n_{p_n} z^{p_n}$$

for $n > 0$.

6.2.1 The optimal web

To illustrate the numbers A_n, w^n_f and w^n_b for a small system of 9 spins, fig. 6.1 shows the *optimal web*: It is a graph whose nodes are all the different optimal states of the system and whose edges are the possible optimal transitions. An optimal path is a path from A_- to A_+ that only passes each n once. Notice the symmetry that all states of a given n are equivalent. Notice also that paths converge in junctions at 'magic' n-values corresponding to filled blocks.

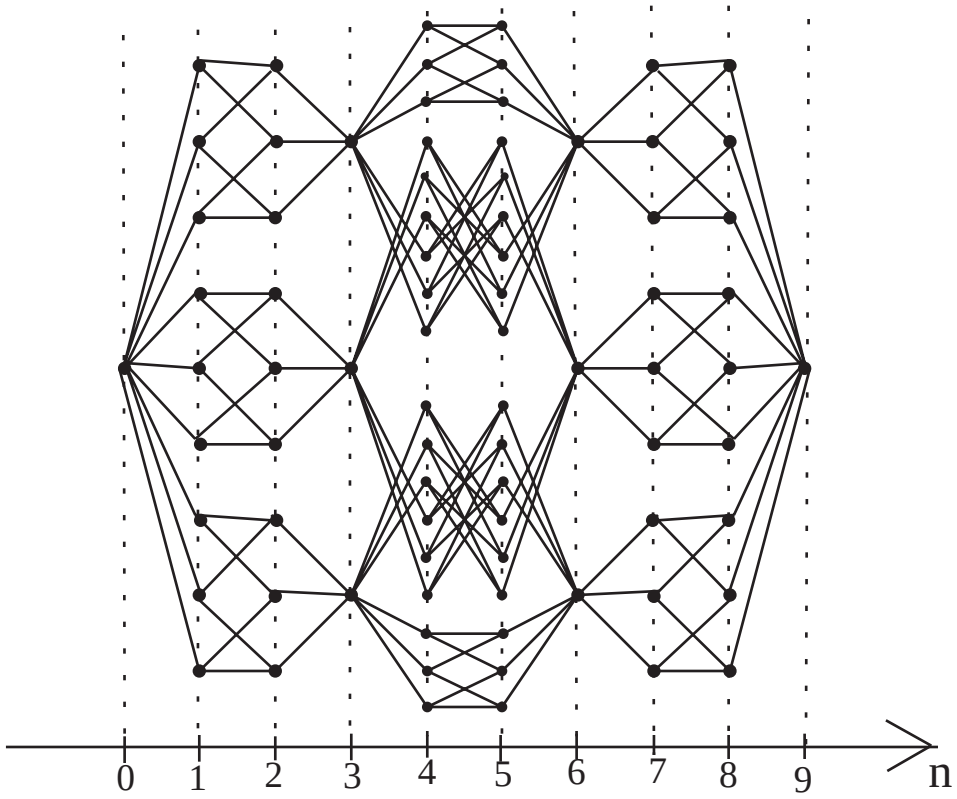


Figure 6.1: The optimal web for $z = 3$ and $L = 2$. Vertices are optimal states and edges are optimal transitions. States are ordered horizontally after n, but there is not assigned any meaning to the vertical axis. Notice that the graph is composed of 3-d cubes in a selfsimilar way (since $z = 3$).

A filled block is locked—it has no internal freedom (cf. filled shells/subshells in atomic structure).

One can imagine (or 'visualize') the energy landscape as a 3-dimensional landscape (but it is not 3-dimensional!) by combining the optimal web graph with the energy profile E_n. The optimal web is the deepest valleys and mountain passes, and in between are all the suboptimal states with higher energies (but there are also valleys in the suboptimal part). It may be that the energy landscape outside the optimal web also has a simple structure that can be calculated exactly, as I have done with the optimal web-part. I think that selfsimilarity dominates the whole energy landscape.

The total number of optimal transitions between n and $n+1$ is $A_n w_f^n = A_{n+1} w_b^{n+1}$. This can be used to write a recursive expression for A_n and we get

$$A_n = A_1 \prod_{i=1}^{n-1} \frac{w_f^i}{w_b^{i+1}} = \frac{z^L}{z^{p_n}} \prod_{i=1}^{n-1} \frac{z - c_{p_i}^i}{c_{p_{i+1}}^{i+1}} \tag{6.6}$$

From this we see how simple the L dependence of A_n is.

6.3 Free energy approximation

What is the free energy at a given (low) temperature and magnetization? In equilibrium most of the probability mass is uniformly distributed among the A_n optimal states with magnetization n that all have the same Gibbs probability, so the entropy is approximately

$$S_n = \ln A_n \tag{6.7}$$

i.e.

$$F_n = E_n - T S_n \tag{6.8}$$

is a free energy approximation. It is most valid at low T. In fig. 6.2 the function F_n is plotted for a range of temperatures. T_c^{Weiss} is the Weiss mean field critical temperature given by eq. 4.11. Notice that the free energy barrier $F_{max} = \max_{0 \le j \le N} F_n$ is lowered with increasing temperature. Notice also that the curve becomes very smooth at a certain temperature ($T = 0.46\, T_c^{Weiss}$). It looks as if the hills and spikes of E_n and of S_n respectively cancel each other near that temperature.

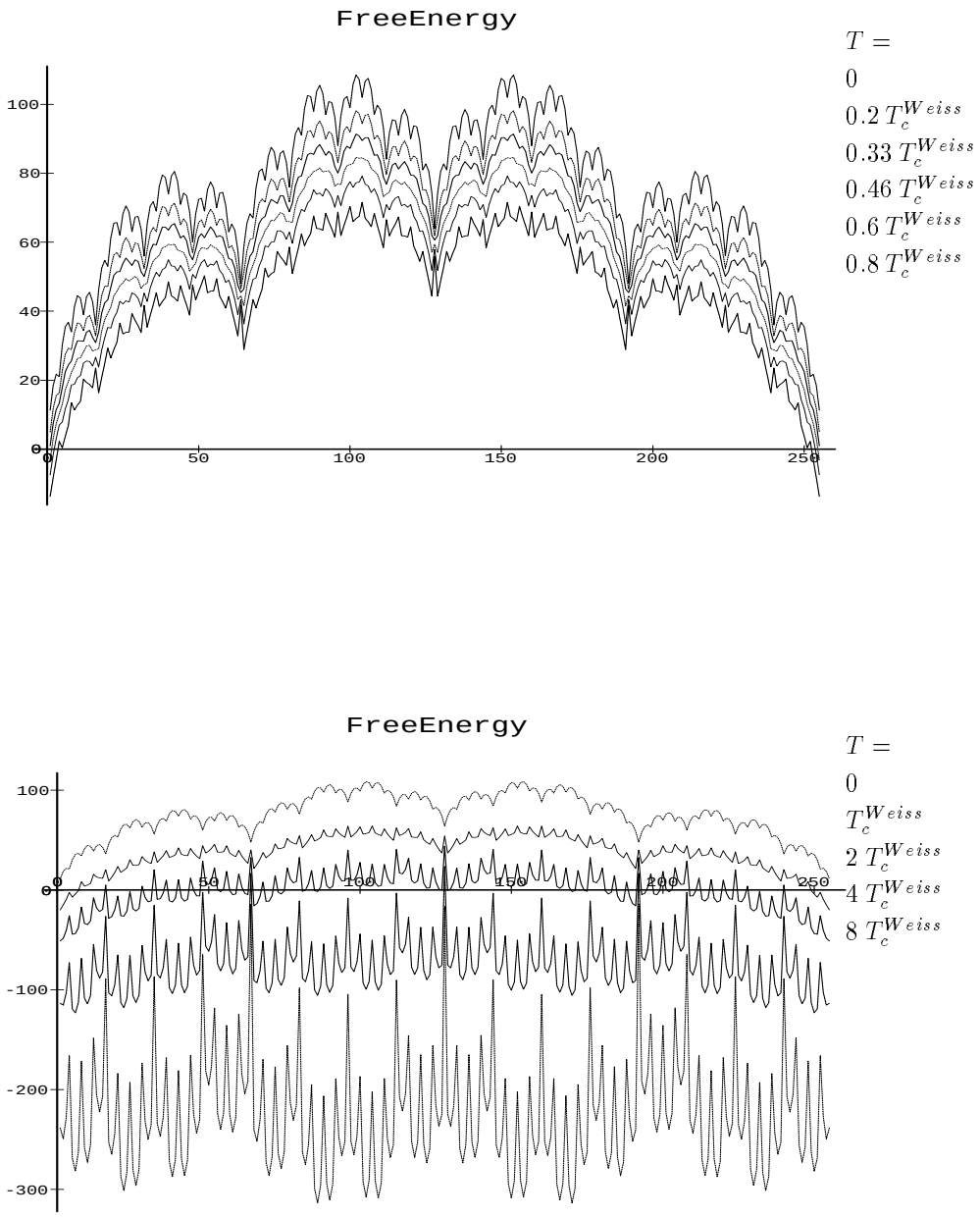


Figure 6.2: F_n (for $0 < n < N$) at a range of temperatures. All other parameters are fixed: $z = 4$, $L = 4$, $\lambda = 1.5$ and $h_0 = 0$. Temperatures are given in units of $T_c^{Weiss} = 5.625$. The energy and entropy is zero for $n = 0$ so $F_0 \equiv 0$.

This pseudo free energy ignores that with increasing temperature the probability mass will move up in the suboptimal part of the energy landscape. Temperatures above the critical temperature are so high that most of the energy landscape is easily accessible. The result is that the system moves towards the zero magnetization ($n = N/2$) region where most of the states are. So for $T > T_c$ a more correct entropy term is

$$S'_n = \ln \left(\begin{array}{c} N \\ n \end{array} \right)$$

In fig. 6.2 it is seen that the entropy term in eq. 6.8 is dominating for $T > T_c^{Weiss}$ where the approximation does not hold. My reason for including these high temperatures in the figure is only to show how the entropy term behaves. Eq. 6.8 is probably a good approximation when the shape of the free energy barriers is dominated by energy and not entropy. Energy should at least win over our first entropy approximation and this is true for $T < 0.5\ T_c^{Weiss}$ in fig. 6.2.

We cannot so easily make more progress analytically with the function F_n. Now we are heading towards numerical work.

6.4 Timescales

The function F_n gives us the relevant free energy barriers between different magnetizations:

$$B_F(n, m) = \max_j (F_j - F_n) \tag{6.9}$$

where j runs between n and m (cf. eq. 5.3). The timescale τ_n before the system explores states n is

$$\tau_n \propto e^{B_F(0,n)/T} \tag{6.10}$$

where $B_F(0, n) = \max_{0 \le j \le n} F_j$. We got the same expression by a heuristic argument in eq. 6.2, which can be seen by rewriting it:

$$\begin{aligned} \tau_n &\propto e^{\max F_j/T} \\ &= \max_{0 \le j \le n} e^{F_j/T} \\ &= \max_{0 \le j \le n} e^{E_j/T} A_j^{-1} \end{aligned}$$

Define *the component* $[0, n(F)]$ to be the states accessible from A_- without having to surmount a free energy barrier higher than F. These states will have magnetizations n where $B_F(0,n) \leq F$ and n will lie in an interval $[0, n(F)]$, see fig. 6.3. The system diffuses back and forth in part a of the

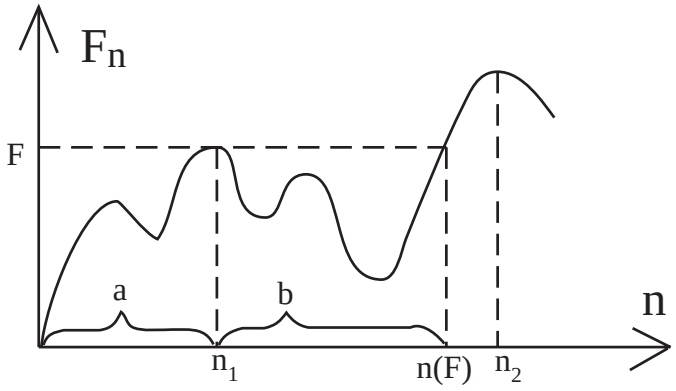


Figure 6.3: Barriers in F_n define components.

component, spending most of the time in the n with lowest free energy, until on timescale $\propto e^{F/T}$ it gets to the barrier top at n_1, and on the same timescale (immediately) falls down in the rest of the component, part b. This is the average behaviour of an ensemble of systems. Part a and b will equilibrate on timescale $e^{F/T}$. The next barrier at n_2 will be crossed much later, on timescale $\propto e^{\beta F_{n_2}}$.

The function F_n has these successively higher barriers as long as the field h_0 (the overall slope) is not too big. So it should be possible to see these equilibrations on different timescales in Monte Carlo simulations. The highest barrier F_{max} of the system determines the nucleation timescale, and after this there are no bigger timescales. We can conclude that if we want F_n to predict these timescales then we must run simulations at low temperature ($T < 0.5\, T_c^{Weiss}$) and small fields h_0.

The *first passage time* τ_n^{fpt} is the first time at which the system in the diffusion process has n spins flipped. I expect that the first passage time τ_n^{fpt} on average will be lower than the equilibration time τ_n, because the system must have visited magnetization n many times before it has equilibrated. But the average first passage time is also expected to be exponential in the

barriers,

$$\tau_n^{fpt} \propto e^{B_F(0,n)/T}, \tag{6.11}$$

the only difference is the prefactor.

Chapter 7

Monte Carlo simulations

In our simulations we want to see the many levels of barriers (and timescales) there are in F_n before the highest barrier F_{max} is reached. This requires that h_0 is small, so that the barriers are high (close to coexistence), and T is small so that the pseudo free energy is a good approximation. But this means that the system will be stuck in the metastable phase for a long time, the nucleation rate will be small. On the other hand we have an experimental timescale on which the phenomena must take place: We only have patience to let computer experiments take hours or one day. We must therefore choose parameters carefully so that the nucleation time does not grow exponentially beyond that limit.

With this time problem in mind I choose Metropolis dynamics which is the fastest (see app. A). I thereby break the close connection to the social hierarchical model there is with Glauber dynamics. But the picture of the dynamics and the predictions about timescales in the previous chapter is independent on this choice of dynamics, after a rescaling of time.

As [9] discusses, the choice between Metropolis and Glauber does not matter in equilibrium studies, but it makes a difference in dynamic studies that people often does not pay attention to. Also details like the choice of candidate for the next state can give results that are not universal but depend on the algorithm.

Because of the time problem I also use a speeded up version of the Metropolis algorithm that I describe in the next section.

7.1 Metropolis speedup

It can be very stupid just to let our computer run the standard Metropolis algorithm (described in section A.6). If we are simulating dynamics at low temperature and there are many metastable states then the system will often be trapped in a metastable state with the probability $e^{-\beta\Delta E/T}$ of escaping it. This exponential can explode if T is low and the metastable state 'deep' enough, making the computer spend hours or weeks running the simulation without changing the configuration of the system.

Instead of repeating the Metropolis step many times without luck, it is smarter to use this idea: Each spin has a probability of being the next spin that will flip. Choose the next spin to flip according to these probabilities and calculate how long time there will go.

In this section I describe my event-driven (one event per step) algorithm which usually is faster than the standard Metropolis algorithm. But it still simulates a system with Metropolis dynamics. It can be used to simulate *any* Ising spin system, so I will describe it without refering to DHM.

7.1.1 Standard

The standard Metropolis algorithm is desribed in general in section A.6. Here I describe it for an Ising model with single-spin-flips. The acceptance probability defined in eq. A.20 is given by eq. B.6. It seems natural that a spin's update rate is independent of the system size, so we let each spin have a constant decision/update rate equal to one. The system's decision rate then is $\lambda' = N$ (section A.3). We will also use continuous time as described in section A.8. This means that we must add $-\frac{1}{N}\ln(x)$ to time in each Metropolis step, where N is the decision rate and x is a random number uniformly distributed in the unit interval[1]. The algorithm thus repeats the following procedure/function.

STANDARD

- Choose a spin i at random (uniformly).
- If $s_i h_i \leq 0$ then FLIP spin i else if $r_1 < e^{-\beta 2 s_i h_i}$ then FLIP i.

[1] I will let the notation $x \in rnd$ have this meaning

- $t := t - \frac{1}{N}\ln(r_2)$

where $r_1, r_2 \in rnd$ and t is time. The program must know the spins s_i and the local fields h_i. Flipping a spin is done by this procedure/function:

FLIP spin i

- $\forall j \neq i: \quad h_j := h_j - 2J_{ij}s_i$ (update local fields)
- change sign of s_i

where we used eq. B.5.

7.1.2 Speedup

For the speedup we need the probability π_j that spin j is the next that flips. With continuous time we have a process like that described in section A.2. What we are asking for is the jump probability given by eq. A.12. Unfortunately I have a notational problem: I use indices i and j in app. A for states and not spins. The primed 'jump' probabilities in eq. A.12 are given by the transition matrix eq. A.22 in which P_{ij} is given by eq. A.21 and P_{accept} is given by eq. B.6. Combining all this we get the probability that spin j flips next

$$\pi_j = \frac{\pi'_j}{\sum_{k=1}^{N} \pi'_k} \tag{7.1}$$

where

$$\pi'_j = \frac{1}{N} \begin{cases} 1 & for \quad s_i h_i \leq 0 \\ e^{-2\beta s_i h_i} & for \quad s_i h_i > 0 \end{cases} \tag{7.2}$$

For the speedup we also need to calculate how long time it takes. The mean waiting time in the state before a spin flip is $1/\lambda$, where λ is the jump rate (not decision rate) of the system, eq. A.11, and equal to the decision rate N times the probability of succes (depends on the state)

$$\lambda = N \sum_{k=1}^{N} \pi'_k \tag{7.3}$$

How do we choose the next flip according to probabilities π_j ? A possible way is to take a random number $r \in rnd$ and find the smallest j for which

$$\sum_{k=1}^{j} \pi_k \geq r \tag{7.4}$$

or equivalently

$$\sum_{k=1}^{j} \pi'_k \geq r \sum_{k=1}^{N} \pi'_k \tag{7.5}$$

If we define some accumulated 'probabilities'

$$a_j \equiv N \sum_{k=1}^{j} \pi'_k \tag{7.6}$$

then we can finally put it all into the following procedure/function:

NEXTFLIP

- For j from 1 to N:
  - If $s_i h_i \leq 0$ then $a_j := 1$, else $a_j := e^{-\beta 2 s_j h_j}$
  - If $j > 1$ then $a_j := a_j + a_{j-1}$
- Find random number $r_1 \in rnd$
- Find smallest j for which $a_j \geq r_1 a_N$ and FLIP that spin.
- $t := t - \frac{1}{a_N} \ln(r_2)$, $r_2 \in rnd$.

where FLIP is the procedure described earlier. The speeded up algorithm simply repeats NEXTFLIP instead of STANDARD. The time step taken by NEXTFLIP (in its last line) can be very large if the system is trapped in a metastable state.

7.1.3 Adaptive algorithm

Using NEXTFLIP once is not always faster than repeating STANDARD until a spin is flipped. The reason is that NEXTFLIP involves more computations than STANDARD. It depends on the energy landscape and the temperature which way is fastest. Is it possible to let the program decide on the fly whether to use NEXTFLIP or repeat STANDARD? It turns out that a simple property of the current system state determines what method is best, with almost no extra computation required.

Let τ_{stand} and τ_{next} be the expected CPU times spent by STANDARD and NEXTFLIP respectively. Let m be the expected number of times STANDARD must be repeated to flip a spin. STANDARD is fastest if $m\tau_{stand} < \tau_{next}$.

Actually we cannot always be sure which way is fastest. In that case it is safest to use NEXTFLIP since it has a big but fixed cost, while repeating STANDARD has an unlimited cost (could be years!). We need some condition to check which if true tells us that STANDARD is fastest and if false can not tell which is fastest in which case we choose NEXTFLIP.

The mean waiting time before a spin flip is

$$\frac{1}{\lambda} = \frac{1}{N\sum_{k=1}^{N}\pi'_k} \tag{7.7}$$

The mean time step taken by STANDARD is $1/N$, so

$$m = \frac{1}{\sum_{k=1}^{N}\pi'_k} \tag{7.8}$$

Define the maximum possible energy increase in the current state

$$wall = \max_{1\le i\le N} 2s_i h_i \tag{7.9}$$

(called wall because a wall tries to prevent you from going in its direction). Then we have

$$\pi'_{min} \equiv \min_{1\le i\le N}\pi'_i = \frac{1}{N}\begin{cases} 1 & for \quad wall \le 0 \\ e^{-\beta wall} & for \quad wall > 0 \end{cases} \tag{7.10}$$

We see then

$$m = \frac{1}{\sum_{k=1}^{N}\pi'_k} \le \frac{1}{N\pi'_{min}} \tag{7.11}$$

We find

$$\begin{aligned} m < \frac{\tau_{next}}{\tau_{stand}} \quad &\Longleftarrow \quad \frac{1}{N\pi'_{min}} < \frac{\tau_{next}}{\tau_{stand}} \\ &\Longleftrightarrow \quad wall < \frac{1}{\beta}\ln\frac{\tau_{next}}{\tau_{stand}} \end{aligned} \tag{7.12}$$

from which we get the last inequality as a sufficient condition that STANDARD is fastest.

I estimate τ_{stand} and τ_{next} by counting the number of times the exponential or logarithmic function is invoked. NEXTFLIP calculates for each spin (i.e. N times)

$$\begin{cases} 1 \\ e^{-\beta\Delta E} \end{cases}$$

and takes the logarithm once, so $\tau_{next} \sim N + 1$. STANDARD calculates

$$\begin{cases} 1 \\ e^{-\beta\Delta E} \end{cases}$$

once and the log once, so $\tau_{stand} \sim 2$. So I estimate[2]

$$\frac{\tau_{next}}{\tau_{stand}} = \frac{N}{2} \tag{7.13}$$

The condition we have been looking for becomes

$$wall < T \ln \frac{N}{2} \equiv threshold \tag{7.14}$$

where $threshold$ is a constant. The adaptive algorithm repeats

QUICKFLIP

- If $wall < threshold$ then STANDARD
- If $wall \geq threshold$ then NEXTFLIP.

Of course $wall$ has to be updated each time the state changes, so FLIP should be

FLIP spin i

- $\forall j \neq i: \quad h_j := h_j - 2J_{ij}s_i$ (update local fields)

[2]I measured the mean τ_{stand} and τ_{next} simulating DHM with $z = 3, L = 5, T = 1, h_0 = 0.1$ and initial configuration given by a droplet of 27 up spins in a uniform background. Running a Turbopascal program I found $\frac{\tau_{next}}{\tau_{stand}} = 180$ and running a C program I found $\frac{\tau_{next}}{\tau_{stand}} = 155$. Compare with the estimate $\frac{N}{2} = 122$.

- $\forall j$: If $2s_jh_j > wall$ then $wall := 2s_jh_j$ (update wall)
- change sign of s_i

We see that the extra computational cost of deciding on the fly which way to use is relatively low.

7.1.4 About the method

The relative speeds of flipping a spin by the two methods depend on the temperature and energy landscape. NEXTFLIP does it in time τ_{next} which does not vary much, and repeating STANDARD does it in time $m\tau_{stand}$, where m can grow exponentially with β. In other words, the speeded up method is exponentially faster than the standard method.

I have described the speedup for Metropolis dynamics but it is just as easy to do it with Glauber dynamics. It requires only a few changes in NEXTFLIP.

A similar speedup algorithm called *the n-fold way* was introduced by Bortz, Kalos and Lebowitz [11]. Instead of NEXTFLIP they study an algorithm designed to simulate ferromagnetic square lattice Ising systems, but the basic idea is the same. They do not have the adaptive algorithm that I use. They find the method to be faster than the standard algorithm at $T < T_c$, but about equally fast at $T = T_c$. Recently [12] has made a generalization of the method to further increase the speed of the n-fold way by many orders of magnitude.

7.2 What to measure?

The relaxation time is a quantity defined for an ensemble of systems. The first passage time is a quantity defined for a single system. Measuring first passage times in simulations is therefore easier to administrate. The relaxation time is also longer than the first passage time so the ensemble of systems in relaxation time experiments must be run for a longer time than the single systems in first passage time experiments. I choose to measure first passage times, hereafter denoted fpt and τ_n.

From eq. 6.11 we have

$$\langle \tau_n \rangle = k e^{B_F(0,n)/T} \tag{7.15}$$

where k is some constant and $\langle \tau_n \rangle$ is the mean fpt. Fortunately, we can find the constant because we can calculate $\langle \tau_1 \rangle$: This is the mean waiting time in which the system is trapped in A_- (because every jump from A_- leads to $n = 1$). The jump rate in A_- is given by eq. 7.3 in which $\pi'_k = 1/N \exp(-\beta E_1)$ for all k, because every spin flip has energy increase E_1. So the escape time in A_- is

$$\tau_1 = \frac{1}{\lambda} = \frac{1}{N e^{-\beta E_1}} = e^{\beta F_1} = e^{B_F(0,1)/T} \tag{7.16}$$

from which we see that $k = 1$.

Measuring T in units of T_c^{Weiss} has an advantage: At h_0 we have $E_1 = 2T_c^{Weiss}$ (see app. B) so

$$\tau_1 = \frac{1}{N} e^{2\frac{T_c^{Weiss}}{T}} \tag{7.17}$$

If T_c^{Weiss}/T is kept constant then the escape time τ_1 is constant when we change λ. So by measuring T in units of T_c^{Weiss} we have in some sense that the behaviour of the dynamics at a certain temperature does not change when we change λ.

In each run of the system I start in the uniform all down state and measure the first time τ_n there are n up spins until $n = N$. In fig. 7.1 I show $\log_{10} \tau_n$ measured for four different runs at the same parameter values. Of course τ_n differs from run to run, so we must average over the runs. In the next section I present the experiments. Each experiment is done at a fixed set of parameters and consists of a number of runs from which I measure $\langle \tau_n \rangle$, $\langle \tau_n^2 \rangle$, $\langle \log_{10} \tau_n \rangle$ and $\langle (\log_{10} \tau_n)^2 \rangle$. I want to use base 10 logarithms instead of ln. For each experiment I present graphs showing $\log_{10} \langle \tau_n \rangle$, $\langle \log_{10} \tau_n \rangle$ and the standard deviation of $\langle \log_{10} \tau_n \rangle$ given by

$$\sigma^2 = \langle (\log_{10} \tau_n)^2 \rangle - \langle \log_{10} \tau_n \rangle^2 \tag{7.18}$$

I try both to take the average before and after I take the log. Logarithms are convex functions so $\langle \log X \rangle \leq \log \langle X \rangle$ for each stochastic variable X. I compare with the prediction eq. 6.11

$$\log_{10} \tau_n = \frac{B_F(0,n)}{T \ln 10} \tag{7.19}$$

by showing graphs of $\frac{F_n}{T \ln 10}$.

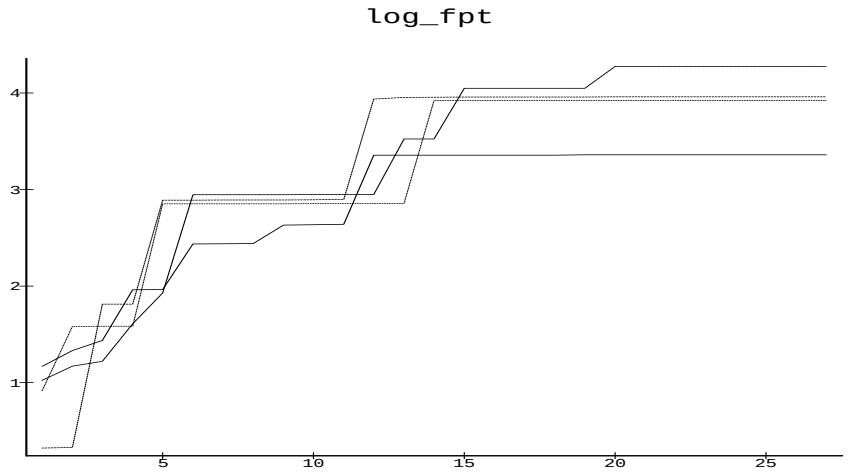


Figure 7.1: $\log_{10}\tau_n$ shown for 4 different runs at $z = 3$, $L = 3$, $\lambda = 1.6$, $h_0 = 0.05$ and $T = 0.4\,T_c^{Weiss}$.

I mostly run small systems because larger systems take longer time to simulate. But they should not be too small because we want to see many levels of barriers. On the other hand the lowest level barriers in a large system may not be visible: At a certain temperature barriers below some crossover size are 'washed away', they are not felt by the dynamics, because T is too high.

7.3 Experiments

Figures 7.3 to 7.5 show the data from a series of experiments run at a range of temperatures for the system with $z = 3$, $L = 3$ (i.e. $N = 27$), $h_0 = 0.05$ and $\lambda = 1.6$. The critical temperature is $T_c^{Weiss} = 3.5697$. The F_n functions for these systems are shown in fig. 7.2. For the lowest temperature we see that the curves for $\log_{10}\langle\tau_n\rangle$ and $\langle\log_{10}\tau_n\rangle$ follows the $\frac{F_n}{T\ln 10}$ curve closely the way we predicted in eq. 7.19. The curves have three plateaus at $2 \le n \le 3$, $7 \le n \le 10$ and $13 \le n$, indicating that a new component is being quickly explored after having overcome a highest barrier, before the next even higher barrier is encountered. This diffusion into a valley takes place on the same

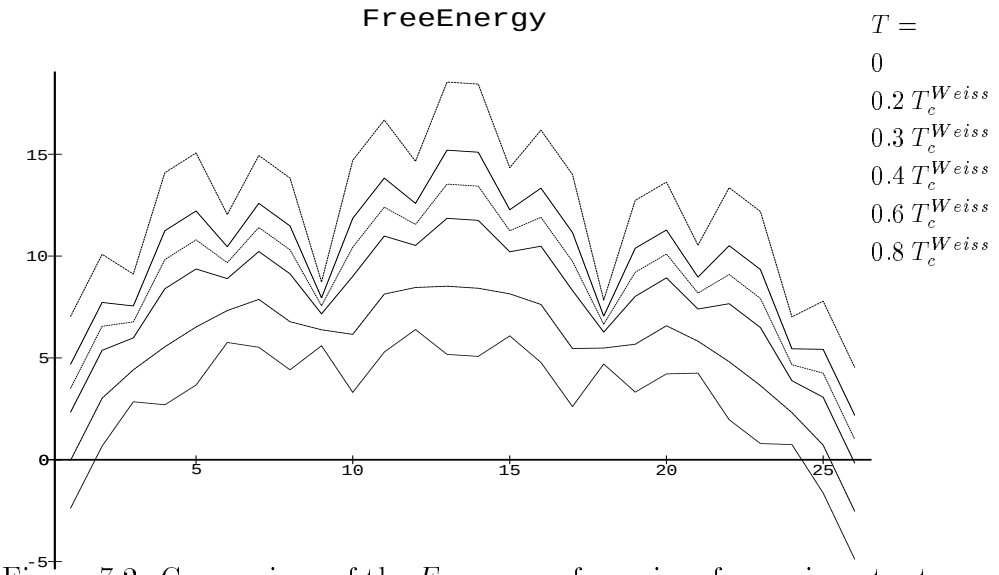


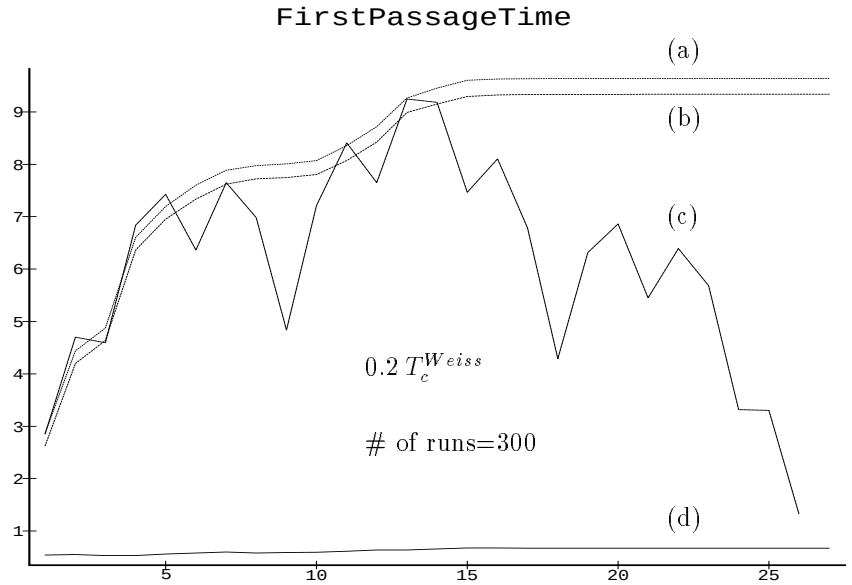
Figure 7.2: Comparison of the F_n-curves of a series of experiments at a range of temperatures and $z = 3$, $L = 3$, $\lambda = 1.6$ and $h_0 = 0.05$.

timescale, but the timescales grow with larger components. The curves are almost constant and equal to the nucleation timescale after the highest barrier F_{max} at $n = 13$, which means that once a critical droplet has appeared, it takes over the whole system on a short timescale. The $\langle \log_{10} \tau_n \rangle$ curves are lower than the $\log_{10} \langle \tau_n \rangle$ curves as expected, and the $\log_{10} \langle \tau_n \rangle$ fits $\frac{F_n}{T \ln 10}$ exactly at $n = 1$ where we could calculate $\langle \tau_1 \rangle$. It is difficult to say which of the two is the 'right' one that fits the landscape best.

As the temperature rises the two time curves still follow the free energy roughly, but they become more smoothed out. At $T = 0.4\,T_c^{Weiss}$ they have almost lost the lower level plateaus but the nucleation time plateau is still there.

For the temperatures $T = 0.6\,T_c^{Weiss}$ and $T = 0.8\,T_c^{Weiss}$ our free energy can no longer be used as prediction. They only roughly predict the nucleation time. At the highest temperature the system seems to have an equilibrium magnetization away from $n = N$, where the system stays until it finds the state $n = N$ on a much larger timescale.

If we look closer at the low temperature curves we see that they still are a bit smoothed out. But we can not expect eq. 7.19 to hold exactly. For example after F_{max} the log-curves do not become truly constant even when



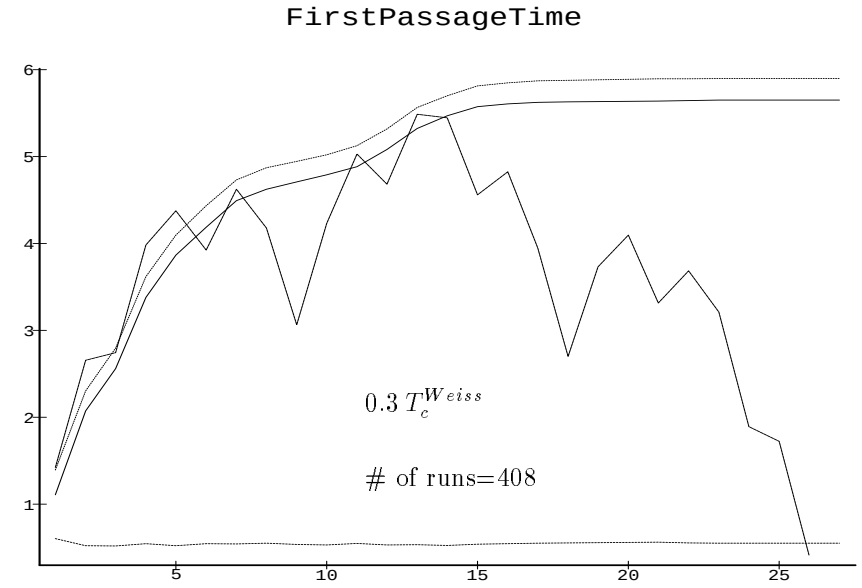


Figure 7.3: Experiments at $0.2T_c^{Weiss}$ and $0.3T_c^{Weiss}$. ($z = 3$, $L = 3$, $h_0 = 0.05$ and $\lambda = 1.6$). Graphs showing (a) $\log_{10}\langle\tau_n\rangle$, (b) $\langle\log_{10}\tau_n\rangle$, (c) $\frac{F_n}{T\ln 10}$ and (d) the standard deviation of $\langle\log_{10}\tau_n\rangle$. The numbers of runs in each experiment are shown.

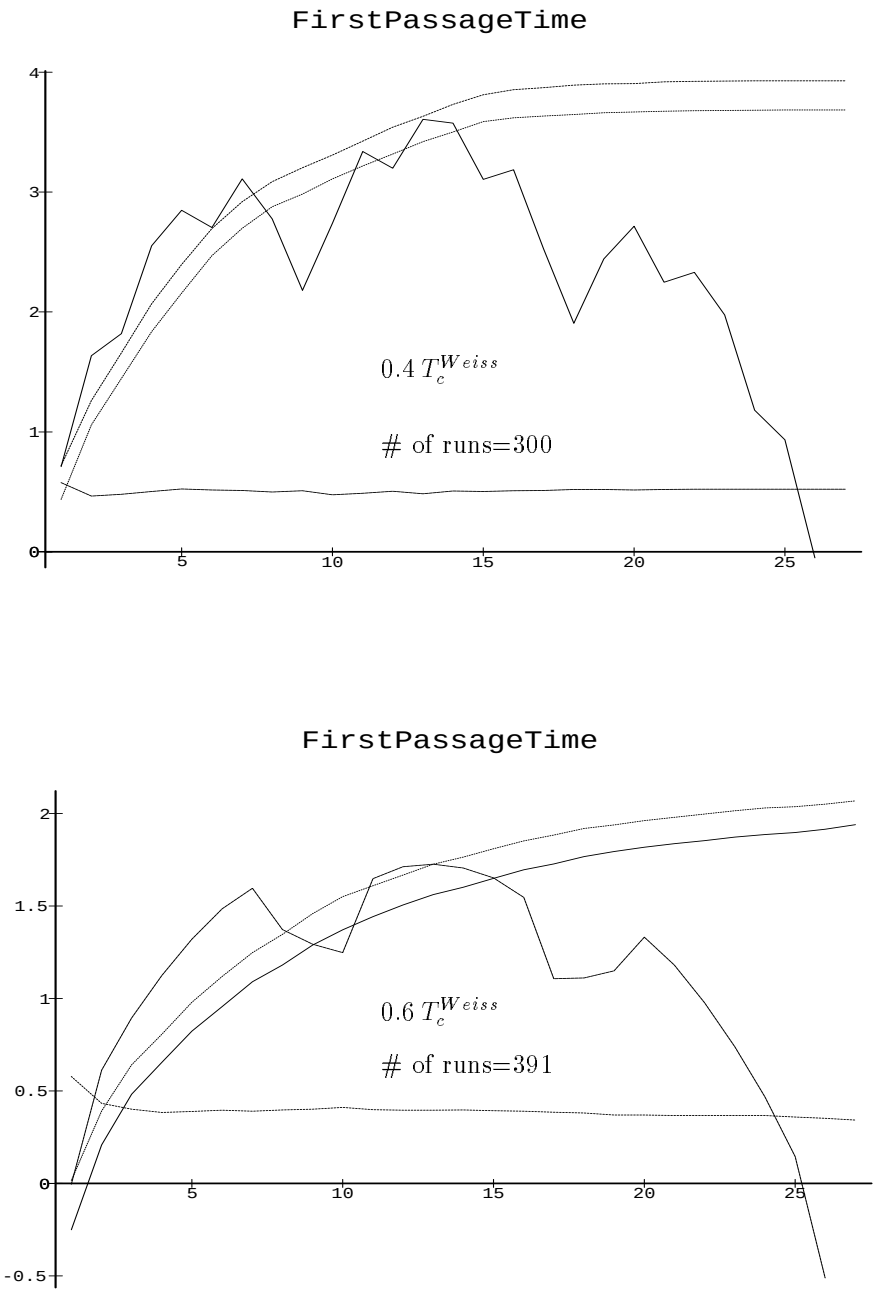


Figure 7.4: Experiments at temperatures $0.4\,T_c^{Weiss}$ and $0.6\,T_c^{Weiss}$. ($z = 3$, $L = 3$, $h_0 = 0.05$ and $\lambda = 1.6$).

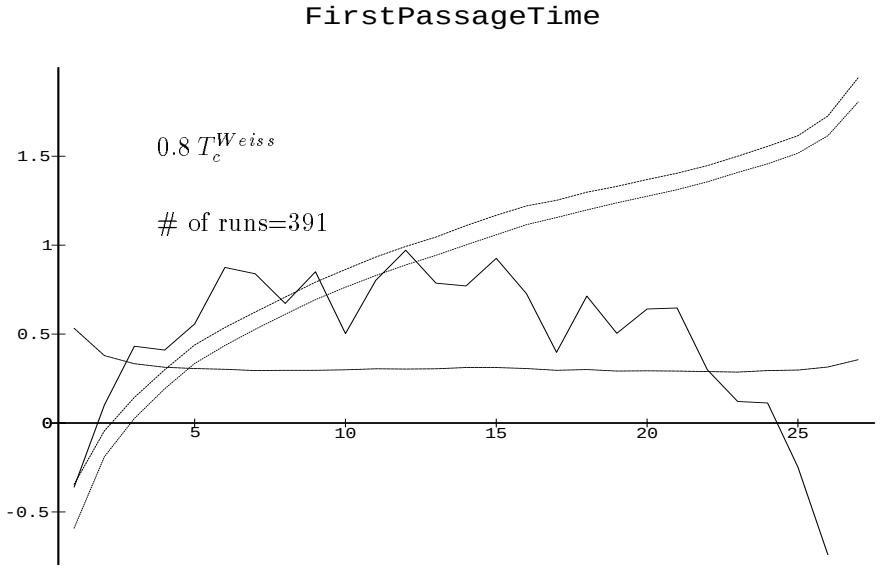


Figure 7.5: Experiment at temperature $0.8\,T_c^{Weiss}$. ($z = 3$, $L = 3$, $h_0 = 0.05$ and $\lambda = 1.6$).

$T \to 0$ because many spinflips do not happen instantaneously. There is also a probability when the system is at a mountain top that it will move back again to lower n-values instead of forward.

7.3.1 Constant standard deviation

We notice two things about the graphs: The standard deviation of $\langle \log_{10} \tau_n \rangle$ and the difference between the two log-curves are almost constant as a function of n. This can be understood if we assume that τ_n is exponentially distributed for each n: We can not see this distribution from our graphs, only averages. But we know that τ_1 is exponentially distributed. It is a reasonable assumption also for $n > 1$. We assume that for any n

$$P(\tau_n > t) = e^{-\lambda_n t} \tag{7.20}$$

with some parameter λ_n. Then we see from

$$P(\lambda_n \tau_n > t) = P(\tau_n > \frac{t}{\lambda_n}) = e^{-t} \tag{7.21}$$

that $x = \lambda_n \tau_n$ is the exponentially distributed variable with mean 1. Now what is the difference between the log of the mean and the mean of the log of τ_n?

$$\ln\langle\tau_n\rangle - \langle\ln\tau_n\rangle = \ln\langle x\rangle - \ln\lambda - (\langle\ln x\rangle - \ln\lambda) = -\langle\ln x\rangle \tag{7.22}$$

because $\langle x\rangle = 1$. We see that the difference does not depend on λ_n or n. It is a constant and equal to

$$-\langle\ln x\rangle = -\int_0^\infty \ln x\, e^{-x} dx = -(-\gamma) \approx 0.577 \tag{7.23}$$

where γ is Euler's constant. We use base 10 logarithms, $\log_{10} = \frac{\ln}{\ln 10}$, for which we find

$$\log_{10}\langle\tau_n\rangle - \langle\log_{10}\tau_n\rangle = \frac{\ln\langle\tau_n\rangle - \langle\ln\tau_n\rangle}{\ln 10} = \frac{\gamma}{\ln 10} \approx 0.25 \tag{7.24}$$

This is close to the 'constant' difference in the figures between the log-mean and mean-log curves.

The standard deviation is

$$\begin{aligned}
\sigma^2(\log_{10}\tau_n) &= \frac{\langle(\ln\tau_n)^2\rangle - \langle\ln\tau_n\rangle^2}{\ln 10} \\
&= \frac{1}{\ln 10}\left(\langle(\ln x - \ln\lambda_n)^2\rangle - \langle\ln x - \ln\lambda_n\rangle^2\right) \\
&= \frac{\langle(\ln x)^2\rangle - \langle\ln x\rangle^2}{\ln 10} = \sigma^2(\log_{10} x)
\end{aligned} \tag{7.25}$$

Thus, it is also a constant and

$$\begin{aligned}
\sigma^2(\ln x) &= \int_0^\infty \ln^2 x\, e^{-x} dx - \left(\int_0^\infty \ln x\, e^{-x} dx\right)^2 \\
&= \pi^2/6 - \gamma - (-\gamma) = \frac{\pi^2}{6}
\end{aligned} \tag{7.26}$$

So $\sigma^2(\log_{10}\tau_n) = \frac{\pi^2/6}{\ln 10} \approx 0.71$. This value is also close to the SD values from the simulation data.

If we know that τ_n is exponentially distributed there will be no information in the graphs of the standard deviation besides the finite-sample fluctuations and therefore no reason to show them. But we do not know the

true distributions, so the graphs do have information. The (almost) constant values lend a certain plausibility to the hypothesis that τ_n is exponentially distributed for any n. I believe that the distributions of τ_n are not very different from exponential distributions.

I also use the SD-curves as a clue telling if the number of runs is sufficiently large. I think I have enough runs when the fluctuations in the SD-curves are small, but I have not tried to increase the number of runs by an order of magnitude to see if the results are unchanged.

For unknown reasons the SD-curves have a somewhat higher value at $n = 1$ than at $n > 1$, especially at higher temperatures.

7.3.2 Lower temperatures

When I go to $T < 0.2\, T_c^{Weiss}$ I soon run into the region where simulations become impossible because timescales grow explosively. Here I only see the system flipping one spin in the state A_- and then immediately flip it back. The Metropolis speedup does not solve the problem of long computation time. It only kills the long waiting time before a single spin flip, but it takes many succesive spin flips to climb the barriers (but see [12]).

I push the temperature limit by having small systems ($N = 9$) and either a large λ (low energy barriers) in fig. 7.6, or a low λ but instead a large h_0 in fig. 7.7. Fig. 7.6 does not tell us anything new, the log-time curves fit the prediction nicely, but in fig. 7.7 we notice that at $n = 2$ the log-time curves are below the prediction. The system reaches $n = 2$ quicker than predicted. Is it because the pathways of the dynamical evolution can make 'shortcuts' in the suboptimal part of the energy landscape? (The actual pathways of the dynamics are mostly in the suboptimal part of the energy landscape: Mostly we see a droplet growing in a *non-uniform* background. But the energy curves for droplet growth in a nearly uniform background are very similar to the optimal droplet energy curves. I think this is the reason that the optimal paths can predict dynamics. When I talk about shortcuts I mean parts of the suboptimal energy landscape that do not 'support' the predictions of optimal paths.)

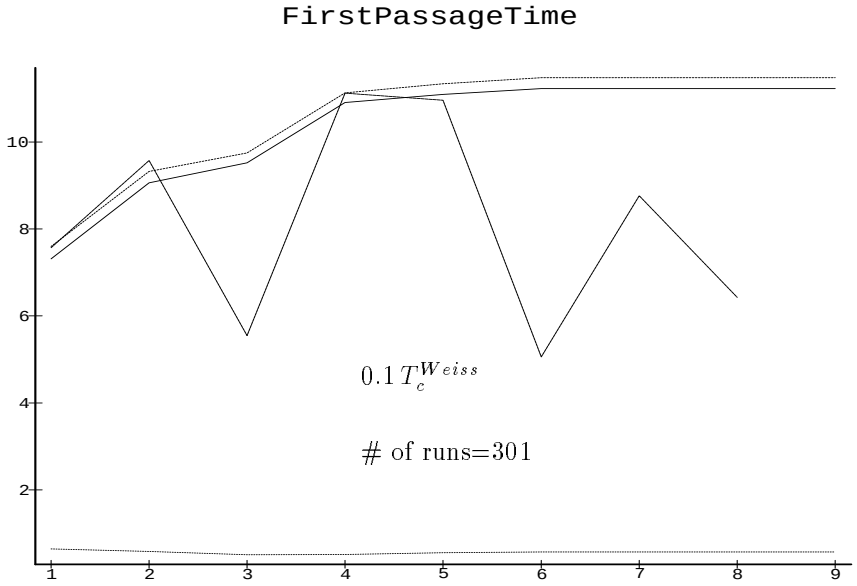


Figure 7.6: Experiment at $T = 0.1\,T_c^{Weiss}$, $z = 3$, $L = 2$, $h_0 = 0.05$ and $\lambda = 2$.

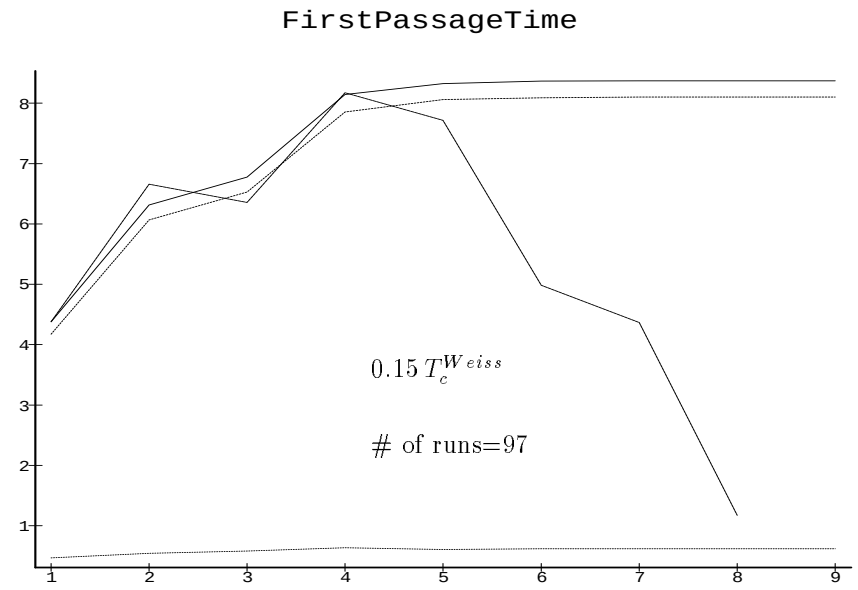
7.3.3 Other experiments

Fig. 7.8 show some runs done with a negative h_0. This means that the system starts in its stable phase and we run it until it finds the metastable phase. As we see the F_n function still predicts the timescales of the first passage time.

In fig. 7.9 I have a large system ($N = 64$) with zero field so that there are two symmetrical 'twin peaks' of height F_{max}. We see that the second peak raises the plateau a little, even though the peak is not higher than the first. Are there one or two components after the first peak?

Fig. 7.10 shows experiments at low λ values. There is not much hierarchical structure at low λ. We see that the nucleation time at $\lambda = 0.6$ is clearly lower than predicted. Is it because of shortcuts in the suboptimal energy landscape? It could be that for $\lambda < 1$ the suboptimal part plays a non-negligible role. I have not investigated this further.

The simulation data open a world of questions that can be studied. With my 10-15 experiments I have only tested the predictions at varying parameters. A study of some of the other details needs more theory to guide simulations systematically.



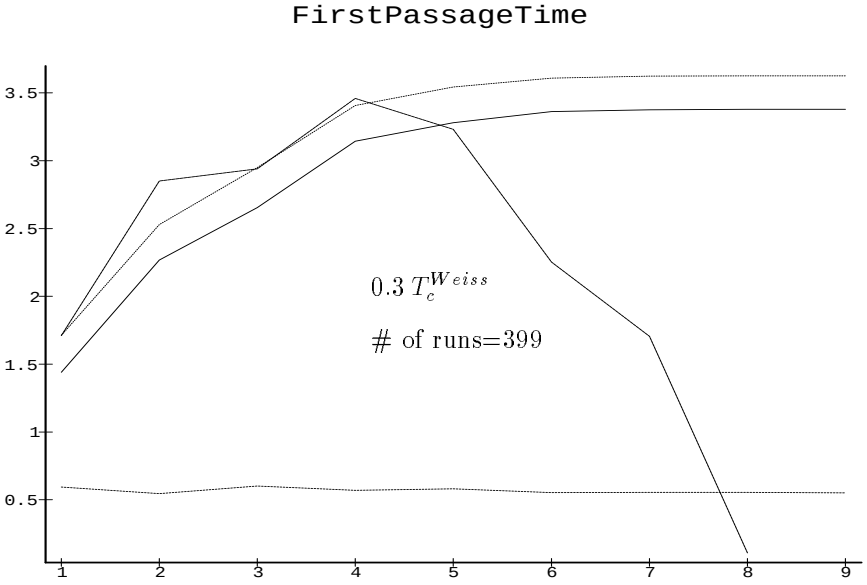


Figure 7.7: Experiments at temperatures $0.15\,T_c^{Weiss}$ and $0.3\,T_c^{Weiss}$. ($z = 3$, $L = 2$, $h_0 = 0.3$ and $\lambda = 1.1$).

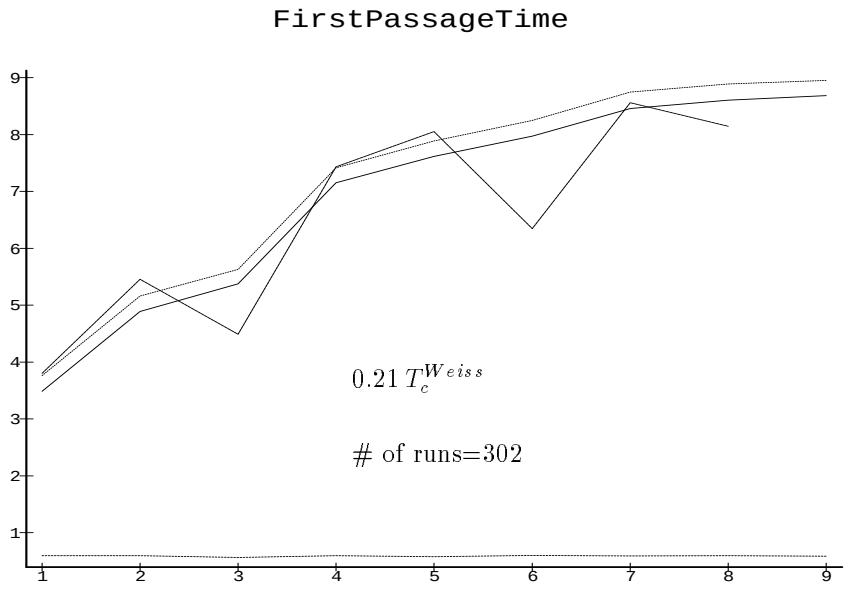


Figure 7.8: Experiment at $T = 0.21\, T_c^{Weiss}$, $z = 3$, $L = 2$, $h_0 = -0.4$ and $\lambda = 2$.

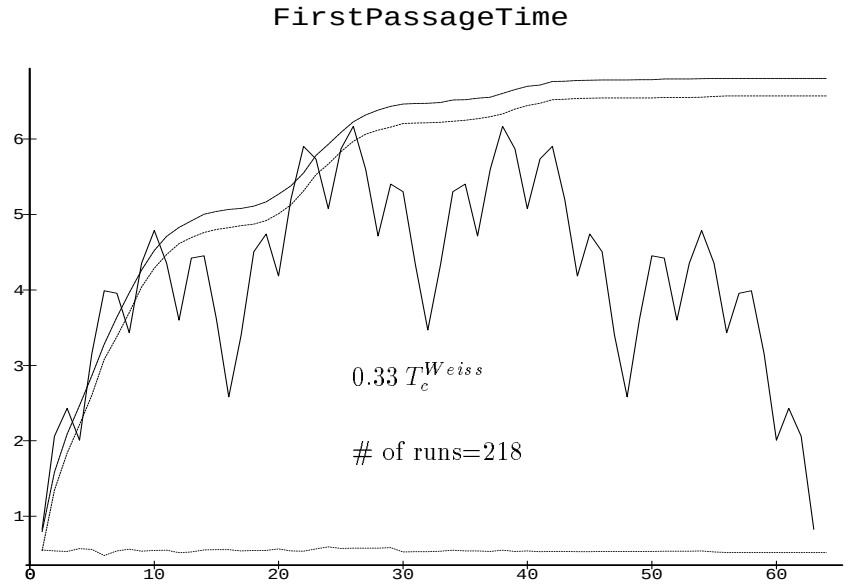


Figure 7.9: Experiment at $T = 0.33 T_c^{Weiss}$, $z = 4$, $L = 3$, $h_0 = 0$ and $\lambda = 1.8$.

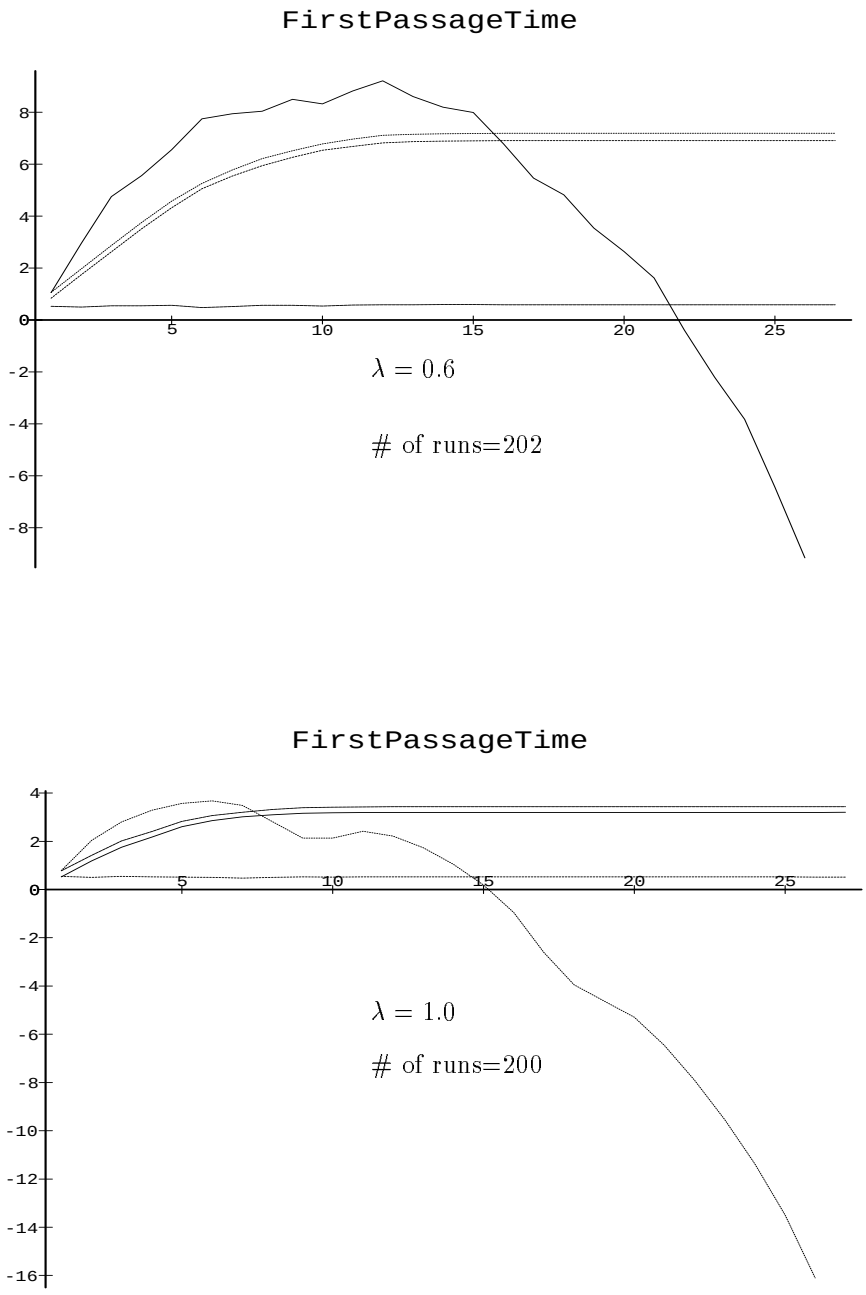


Figure 7.10: Experiments at $\lambda = 0.6$ and $\lambda = 1.0$. ($z = 3$, $L = 3$, $h_0 = 1.4$ and $T = 0.3\, T_c^{Weiss}$).

7.3.4 Number of spin flips

I made my program count the total number of spin flips performed until the state $n = N$ was found. I did not record these numbers in the simulation data. But afterwards I found the numbers in some garbage data for some of the experiments. For the first experiment shown in fig. 7.3 ($T = 0.2 T_c^{Weiss}$) I found that there is an average of 14 million spin flips per run. The experiment in fig. 7.9 has an average of 2.2 million spin flips per run and the experiment in fig. 7.10 ($\lambda = 0.6$) has 2.9 million.

These numbers are interesting when I try to imagine the simulation scenarios in the context of the social hierarchical model in the last chapter.

7.4 Conclusion

The free energy approximation worked. The F_n function was able to catch the interesting physics (plateaus in the log-fpt curves) at low enough T. It was not clear that the approximation would work at the temperatures where we can simulate the system. And it was not clear that a one dimensional picture could be used (F_n is a function of one variable). We now understand nucleation in DHM as intended.

Chapter 8

Discussion and conclusion

8.1 Modelling societies

The social models we have worked with are of a type that have no (or very few) concrete experimental data to predict and be compared with. It is not easy to make measurements on social or economic systems. But maybe the computer networks will open social worlds where all the activity can be observed. It has been a challenge to work with these models that are both physics and social sciences, and to consider their 'goodness' as models.

In this kind of modelling we do not know the details. With uncertain knowledge we can make a very simplified and coarse grained model, but we can also introduce fine grained details and knowledge with our assumptions, if we remember to remove them again. We did both, but I will discuss the latter. If we introduce details which we do not know and that could be different then we must know how much our results depend on them. How stable are our results to perturbations in the assumptions and alternatives? The results should not depend on those details, they should be universal. We choose a road when we construct and derive our model, but many roads should lead to the same result. The details that we can not know should therefore not be contained in the results, they must disappear on the way. This corresponds to having many decimals in intermediate calculations before truncating them in the final result. This is what we saw when we derived the threshold rule eq. 2.19, where all the details about agents' expectations had vanished. And we found a threshold rule to be the result for any expectation

function in the bandwagon class. The more our result depend on the details, the less likely it is to find the same thing empirically. We should not delve into special technical questions and loose the general view. One example of this is the fractal properties of the energy landscape discussed in section 5.3.4: Studying them would probably lead into a dead end.

It is not so important what the details are if they will disappear. Therefore we can be sloppy, make approximations, make unreasonable and unrealistic assumptions and have inconsistencies if we know they do not change the end result. We must know the overall course—what we try to model. For example we may want to model a class of phenomena where it is a threshold rule (Ising spin like interaction), and not the rule we get with opportunistic expectation functions. It is these qualitative differences that should be discussed and not details.

Even when details are not important we usually derive things mathematically correct, and take the logically right steps, for example in the reasoning leading to eq. 2.9. We do it to prevent making errors that will matter in the end. But it is deceptive and can be confusing because we do not really *derive* —we only make our model plausible by taking a possible exact road to it.

Our social models should be judged in this light. We do not know how irrational agents make their decisions in the real world. The model is only made to show that the overall picture *can* be generated by some plausible mechanisms.

A discussion in this philosophy of science spirit is required in principle for any (mathematical) model. With uncertain knowledge it is necessary to think in 'ensembles' of theories/models. Maybe this discussion becomes more important as we go from simple systems in physics to more complex systems in sociological/game theoretical modelling where more details are unknown. Models in physics often need not be discussed because they build on accepted theories or measured facts and because a physicist know what a physical theory is. But a physicist looking at models of people for example do not have that background experience to rely on.

There is also an ethical reason for discussing the limitations of the model: Huberman *et al.* give normative prescriptions of how to structure organizations. Decision makers may uncritically take the results and use in situations where they do not apply. The mathematics behind may give the impression of a quantitative knowledge that we do not have. (But this is still better than astrology etc.)

8.2 Where do our scenarios apply?

I will try to put our results about the dynamics of DHM in relation to the social models.

I have studied dynamical scenarios with many timescales: The hierarchical system changes its overall configuration (level of cooperation) in a nucleation process in which successively higher barriers must be overcome. Loosely speaking, blocks on different levels nucleate on different timescales. When a critical size of the system has changed (the critical droplet) the rest of the system changes on a short timescales. But this is also seen on lower levels: After a part (a block?) of the system has nucleated the rest of the part follows quickly. This gives rise to plateaus in the log-fpt curves. These scenarios were seen when T and h_0 was low.

If T is high then the hierarchical structure does not influence dynamics (at least not the log-fpt curves) and a hierarchical model is unnecessary. If T is low but h_0 is high we get another kind of scenarios. High and low h_0 scenarios[1] are illustrated in fig. 8.1. In high h_0 scenarios the barrier F_{max} is low and the critical droplet small. There are not many timescales. The interesting physics in these scenarios is not the nucleation process, but what happens *after* a critical droplet has appeared (domain growth): Then we have a fast slide downhill the F_n-curve with short stays in the local minima. The magnetization as a function of time has a stepwise increase. It is this behaviour of the social hierarchical model (the cascading phenomenon) that originally interested us in section 3.4. Here the hierarchical structure does not reveal itself in the log-fpt curves, but it appears in the short stays in the valleys of the free energy. I believe the F_n function can predict how long time it spends in these valleys. But instead we have been guided by the idea of many timescales.

Many timescales means that the largest timescale is big: It takes a very long time to nucleate. I found these processes to involve millions of spinflips per run. This means $10^5 - 10^6$ flips per individual (spin). The number of Metropolis *decisions* is probably a number of magnitudes higher and with Glauber it is even higher. How likely is it to find scenarios in real situations in which each individual makes millions of decisions? A decision is a

[1] Known as nucleation near the spinodal and nucleation close to coexistense respectively [10].

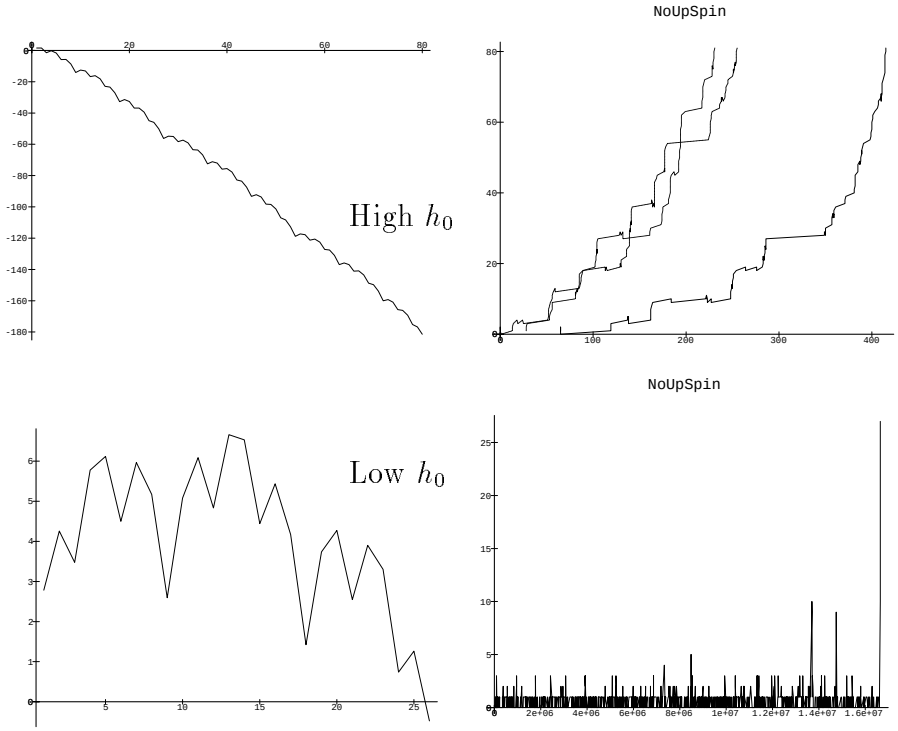


Figure 8.1: Three runs at High h_0 ($h_0 = 0.4\, h_{max}$, $T = 0.15\, T_c^{Weiss}$, $\lambda = 1.8$, $z = 3$ and $L = 4$) and one run at Low h_0 ($h_0 = 0.03\, h_{max}$, $T = 0.2\, T_c^{Weiss}$, $\lambda = 1.8$, $z = 3$ and $L = 3$). Graphs on the left show $\frac{F_n}{T \ln 10}$ and graphs on the right show the number of up spins as function of time.

reconsideration of the individual's own level of cooperation. It involves an assessment of the overall level of cooperation, a decision based on that, and possibly a change of action which must be perceived by the others without much delay compared to the rate of decisions (because I have no delay in the Glauber dynamics). This puts a limit on the rate of decisions of an agent. Typically the agent (for example a human) have other things to think about too, other than deciding whether to continue cooperating or not. With a rate of 100 decisions per day it takes 27 years to make a million decisions. Not many sociological 'games' with such a high rate of decisions have such a long time span. We would not expect the environment, for example the hierarchical structure to be constant in such a long time [Fluidity]. This makes it very restricted where our many timescales scenarios apply. But it is not impossible to find examples. Computers in a network can make these decisions at a much higher rate than humans. With a one second update rate per computer it takes 28 hours to make 10^5 updates per computer. [Behavior, page 82] talks about a millisecond timescale for computational processes to choose strategy vs at least seconds in human societies.

It is much more likely to find examples of the high h_0 scenarios which takes place on shorter timescales, so these scenarios are more interesting in the sociological context, and can be subject of further work.

8.3 Robustness of the model

I only looked at the dynamics of DHM, which is a special case of the social hierarchical model (SHM), with a homogeneously branching rate in the hierarchy tree. The exact analytical 'shape' of different quantities in the model should be different if we try to match a particular social dilemma situation of a hierarchically clustered system. With the discussion in the beginning of this chapter in mind we should consider how robust our results are to such alternatives.

If a social system has Ising spin like interactions (as the threshold rule eq. 3.7) and it is hierarchically clustered, then I expect the hierarchical structure to dominate the energy landscape as we saw for DHM. And then we will be able to see plateaus in the number of cooperators as a function of time or in the log-fpt curves. So I expect the general picture of the two scenarios discussed in the previous section to be very robust to variations in the model.

Our results about the effects of hierarchy on dynamics has therefore some universality.

If we want to estimate *how* robust the results are we can begin with playing with the parameters of either DHM (z, L, h_0, J, T, λ) or SHM (a, tree structure, b, c, H, α). Instead of constants we can have site dependent or time varying quantities (different agents have different thresholds etc.) There are some limits on how much inhomogeneity, diversity and disorder we can introduce before our results break down. I have not estimated these limits. But one of the first things that happens, with only small variations in the parameters, is probably that the degeneracy of optimal paths disappear. This does not change our results. With stronger perturbations we may see optimal paths involving two or more droplets. With enough disorder the energy landscape becomes spin glass like.

I understand the physical and thermodynamical meaning of the parameters of DHM much better than the parameters of SHM. I made a mapping between them for the flat model to transfer this understanding from one to the other, but I did not make the mapping for the hierarchical case. This is necessary if we want to analyse what our simulation results means in the social models. This translation will be more complicated than for the flat model. We saw that the external field becomes non uniform when the tree does not have a constant branching rate z. Such irregular behaviour of the parameters may also make our scenario pictures less stable to variations in the parameters of SHM than in the DHM parameters.

These matters could also be the subject of further work.

I want to separate from the rest of the thesis some theory that can be looked up (at least in principle) in textbooks, both to avoid polluting the thesis with less relevant derivations and to discriminate what is general and what is specific to my problem. But I do not want to refer to the literature all the time, so therefore I have these appendices. I think this is a good way to organize my knowledge. I have made them self-contained texts i.e. not just mathematical passages of the thesis moved to these pages. I use them as a toolbox. Here is the references: [van Kampen, Markovkæder, Critical, Neural, Glasses, Annealing]

Appendix A

Stochastic dynamics

A.1 Master equation

Let me introduce the notation. Consider a homogeneous Markov process in continuous time with a discrete finite set I of states. Let X_t denote the state at time t. $P(X_t = i) = p_i(t)$ is the probability distribution of states at time t. The *transition probability* is the conditional probability $P(X_t = j|X_s = i)$ (called $P_{1|1}$ in [van Kampen]). *Homogeneous* means that the transition probabilities are invariant under time translations so that

$$P(X_{t+h} = j|X_{s+h} = i) = P(X_t = j|X_s = i) \tag{A.1}$$

Thus the transition probability depends on s and t only through the time difference $t - s = \tau$ and we write it as a matrix

$$T_{ij}(\tau) = P(X_{s+\tau} = j|X_s = i) \tag{A.2}$$

(called $T_\tau(j|i)$ in [van Kampen]). The *initial distribution* $p(0)$ determines $p(t)$ by

$$p_j(t) = \sum_i T_{ij}(t)p_i(0) \tag{A.3}$$

or in matrix notation $p(t) = p(0)T(t)$ (but this is not the master equation). The process starts at $t = 0$ in a state Y_0 where it stays in a time interval until t_1. Then it jumps to another state Y_1 $(= X_{t_1})$ where it stays until at t_2 it jumps to another state Y_2 $(= X_{t_2})$ and so on. The *waiting time* $t_{n+1} - t_n$

in state $Y_n = i$ is exponentially distributed with parameter $\lambda_i \geq 0$ so that $P(t_{n+1} - t_n > t) = e^{-\lambda_i t}$. The mean waiting time in state i is $1/\lambda_i$. The rate of jumps in state i is λ_i. The *jump probability* $\pi_{ij} \geq 0$ is the probability in state i that the next state it jumps to is j. We have $\pi_{ii} = 0$ and $\sum_j \pi_{ij} = 1$. If $\lambda_i = 0$ it means that the process stays in i forever once it has gotten there (i is *absorbing*)and then π_{ij} is arbitrary but still restricted by $\sum_j \pi_{ij} = 1$. The *intensity matrix* W is

$$\begin{array}{l} W_{ij} = \lambda_i \pi_{ij} \quad for \quad i \neq j \\ W_{ii} = -\lambda_i \end{array} \tag{A.4}$$

We have $\sum_j W_{ij} = 0$, $W_{ij} \geq 0$ for $i \neq j$ and $W_{ii} \leq 0$. (Van Kampen has *the $\mathcal{W}$-matrix* defined as the matrix W^t ($W_{ij} = \mathcal{W}_{ji}$). Some define W by the following

$$\begin{array}{c} T_{ij}(h) = P(X_{t+h} = j | X_t = i) = W_{ij} h + o(h) = \lambda_i \pi_{ij} h + o(h) \\ T_{ii}(h) = P(X_{t+h} = i | X_t = i) = 1 + W_{ii} h + o(h) = 1 - \lambda_i h + o(h) \end{array} \tag{A.5}$$

so that for $i \neq j$ W_{ij} is the *transition probability per unit time* from i to j. One gets the 'backward' and 'forward' Feller-Kolmogorov equations

$$\tfrac{d}{dt} T(t) \quad = \quad T(t) W \quad = \quad W T(t) \tag{A.6}$$

and differentiating $p(0)T(t)$ gives us the *master equation*

$$\frac{d}{dt} p(t) = p(t) W \tag{A.7}$$

or

$$\begin{aligned} \frac{d}{dt} p_i(t) &= W_{ii} p_i(t) + \sum_{j \neq i} p_j(t) W_{ji} \\ &= (\sum_{j \neq i} p_j(t) W_{ji}) - (\sum_{j \neq i} p_i(t) W_{ij}) \\ &= \sum_{j \neq i} (p_j(t) \lambda_j \pi_{ji} - p_i(t) \lambda_i \pi_{ij}) \end{aligned} \tag{A.8}$$

because $W_{ii} = -\sum_{j \neq i} W_{ij}$. We can also sum over I instead in the last two expressions. For $i \neq j$ the intensity W_{ij} is the fraction of the probability

mass in i that flows to j per unit time. The first sum in the second line is what flows into i and the second is what flows out. The unique solution (when the number of states is finite) to this simple differential equation is

$$p(t) = p(0)e^{tW} \tag{A.9}$$

where the matrix is

$$e^{tW} = \sum_{n=0}^{\infty} \frac{(tW)^n}{n!}$$

Very often λ_i is the same for all i and $1/\lambda_i$ is a characteristic microscopic time τ_0 for the process and the master equation is written

$$\tau_0 \frac{dp_i(t)}{dt} = \sum_j p_j(t)\pi_{ji} - p_i(t)\pi_{ij}$$

A.2 Jumps to the same state

Sometimes the dynamics of the process is constructed with 'jump probabilities' π'_{ij} where $\pi'_{ii} \geq 0$, so that it is possible to jump to i from i (which means not to jump), and with a rate λ'_i of *decisions* or *updates* or *attempts* in state i (the *wake up rate*) so that $P(D_i > t) = e^{-\lambda'_i t}$ where D_i is the time in i before a decision; jump to another state or the same state (no jump). Then what are the jump rates λ_i, the jump probabilities π_{ij} (with $\pi_{ii} = 0$) and the intesity matrix W ? Let J_i be the time spent in i before a jump and let u be the number of jump attempts in a time interval $[0, t]$. The probability of a jump attempt to fail in state i is π'_{ii}. Then

$$\begin{aligned}
P(J_i > t) &= \textstyle\sum_{n=0}^{\infty} P(u = n, \ \ all \ \ attempts \ \ failed) \\
&= \textstyle\sum_{n=0}^{\infty} P(u = n)(\pi'_{ii})^n \\
&= \textstyle\sum_{n=0}^{\infty} \{\frac{(\lambda'_i t)^n}{n!} e^{-\lambda'_i t}\}(\pi'_{ii})^n \qquad\qquad (poisson) \\
&= e^{-\lambda'_i t} \textstyle\sum_{n=0}^{\infty} \frac{(\lambda'_i \pi'_{ii} t)^n}{n!} \\
&= e^{-\lambda'_i (1-\pi'_{ii})t}
\end{aligned} \tag{A.10}$$

So the jump rate is

$$\lambda_i = \lambda'_i(1 - \pi'_{ii}) = \lambda'_i \sum_{k \neq i} \pi'_{ik} \tag{A.11}$$

which means that the rate of successes =rate of attempts times the probability of succes in an attempt. For $i \neq j$ the jump probabilities are

$$\pi_{ij} = \frac{\pi'_{ij}}{\sum_{k \neq i} \pi'_{ik}} = \frac{\pi'_{ij}}{1 - \pi'_{ii}} \tag{A.12}$$

(We use these results in speeding up the metropolis algorithm). The intensity matrix is

$$W_{ij} = \begin{cases} \lambda_i \pi_{ij} = \lambda'_i (1 - \pi'_{ii}) \frac{\pi'_{ij}}{1 - \pi'_{ii}} & for \quad i \neq j \\ -\lambda_i = -\lambda'_i (1 - \pi'_{ii}) & for \quad i = j \end{cases} \tag{A.13}$$

So if we write the master equation with primed quantities instead of unprimed we see that it has the same form:

$$\frac{dp_i(t)}{dt} = \sum_{j \neq i} (p_j(t) \lambda'_j \pi'_{ji} - p_i(t) \lambda'_i \pi'_{ij})$$

A.3 Many parts

Sometimes a Markov process is constructed to model a system consisting of many parts like spins or agents. If I^k is the set of states of part k then $I^1 \times \cdots \times I^N$ is the set of states of the system (N is the number of parts). Each part k makes jumps (changes of its state) spontaneously and there is a waiting time τ^k between jumps of part k. The waiting times of different parts are independent stochastic variables given the state of the system and $P(\tau^k > t) = e^{-\lambda_i^k t}$ where λ_i^k is the rate of jumps of part k in system state i. What about the waiting time τ in system state i before the system makes a jump?

$$\begin{aligned} P(\tau > t) &= P(\tau^1 > t \wedge \cdots \wedge \tau^N > t) &= P(\tau^1 > t) \cdots P(\tau^N > t) \\ & &= e^{-\lambda_i^1 t} \cdots e^{-\lambda_i^N t} \end{aligned} \tag{A.14}$$

The rate of jumps of the system in state i is thus

$$\lambda_i = \sum_k \lambda_i^k \tag{A.15}$$

If for example each part jumps at a constant rate $\lambda_i^k = \alpha$ then the jump rate of the system is $N\alpha$. For each part k there is a jump probability matrix π^k

such that π_{ij}^k is the probability in system state i that the next state is j given that it is part k that changes its state. The jump probabilities of the system is then given by

$$\pi_{ij} = p_i^1 \pi_{ij}^1 + \ldots + p_i^N \pi_{ij}^N$$

where p_i^k is the probability in system state i that it is part k that jumps next.

$$\begin{aligned} p_i^k &= \int_0^\infty \lambda_i^k e^{-\lambda_i^k t} \prod_{l \neq k} P(\tau^l > t) dt \\ &= \lambda_i^k \int_0^\infty e^{-\lambda_i^1 t} \cdots e^{-\lambda_i^N t} dt \\ &= \frac{\lambda_i^k}{\lambda_i^1 + \cdots + \lambda_i^N} = \frac{\lambda_i^k}{\lambda_i} \end{aligned} \tag{A.16}$$

A.4 Stationary states and detailed balance

A *stationary* distribution a is a constant solution to the master equation, in other words a distribution fulfilling

$$\forall t : a = aT(t)$$

or

$$aW = 0$$

saying that a is left eigenvector of $T(t)$ and W, or

$$\forall i : \ \textstyle\sum_{j \neq i} a_j W_{ji} = \sum_{j \neq i} a_i W_{ij}$$

saying that there is a balance between flows 'in' and 'out' for each state which of course makes it constant in time (an equilibrium). A distribution a is stationary if there is *detailed balance* which is the condition

$$\forall i, j : \ a_j W_{ji} = a_i W_{ij} \tag{A.17}$$

but this condition is not necessary for stationarity.

A Markov process is *irreducible* if $\lambda_i > 0$ for all i and if some power π^n of the jump probability matrix has all components positive, so that from any state it is possible to reach any other state. When the process is irreducible no states are *transient* which means that the process will return to any state in finite time (ergodicity). When the Markov process is irreducible there exists exactly one stationary distribution a and $p(t) \to a$ for $t \to \infty$ for all initial distributions $p(0)$. This stationary distribution will have all $a_i > 0$.

A.5 Energy

At this point we imagine that every state i has an *energy* E_i. The *Gibbs distribution* is an assignment of probabilities to states given their energies and a temperature $1/\beta$

$$p_i^{Gibbs} = \frac{1}{Z} e^{-\beta E_i} \tag{A.18}$$

where $Z = \sum_i e^{-\beta E_i}$. Given a probability distribution p_i, a temperature $1/\beta$ and a positive constant Z it is also possible to assign energies E_i so that the distribution is the Gibbs distribution, by

$$\beta E_i = -\ln p_i - \ln Z \tag{A.19}$$

Changing Z just adds a constant to the energy function and changing β just scales the energy with a factor.

A.6 Metropolis algorithm

I will describe different kinds of stochastic Monte Carlo dynamics where the distribution of states will tend to a Gibbs distribution. So I look at a finite number of states, with energies E_i. Each state i also has a set $N(i)$ of states called its *neighbours*. A state i is not its own neighbour: i is not in $N(i)$. The first kind of dynamics is defined by the *standard Metropolis algorithm*. The algorithm is to repeat the following:

- Choose a neighbour $j \in N(i)$ at random to the current state i. Each state in $N(i)$ has uniform probability $= 1/|N(i)|$ of being chosen.
- If $\Delta E = E_j - E_i \leq 0$ then the next state is j.
- If $\Delta E > 0$ then the next state will be cosen at random to be j only with probability $e^{-\beta \Delta E}$, and to be i otherwise.

($T = 1/\beta$ is a temperature). This is the *Metropolis step* and time is increased by 1 each time the step is performed (the time unit is a Metropolis step). The chosen candidate j for the next state in a step is thus accepted with probability

$$P_{accept}(\Delta E) = \begin{cases} 1 & for \quad \Delta E \leq 0 \\ e^{-\beta \Delta E} & for \quad \Delta E > 0 \end{cases} \tag{A.20}$$

The algorithm defines a Markov process in *discrete time*, so let us find the *transition matrix* $\tilde{\Pi}$. First define a matrix that tells us what states are neighbours and what their probabilities are of being chosen: The *proposal distribution* P.

$$P_{ij} = \begin{cases} \frac{1}{|N(i)|} & for \quad j \in N(i) \\ 0 & otherwise \end{cases} \tag{A.21}$$

P is a stochastic matrix: All rows sum up to one. The transition matrix is then

$$\tilde{\Pi}_{ij} = \begin{cases} P_{ij} P_{accept}(\Delta E) & for \quad j \neq i \\ 1 - \sum_{j \neq i} \tilde{\Pi}_{ij} & for \quad i = j \end{cases} \tag{A.22}$$

The rows of this matrix also sum to one, by construction. We want the Gibbs distribution p^{Gibbs} to be a stationary distribution i.e.

$$p^{Gibbs} = p^{Gibbs} \tilde{\Pi}$$

and we obtain this by having detailed balance:

$$\forall i, j : p_i^{Gibbs} \tilde{\Pi}_{ij} = p_j^{Gibbs} \tilde{\Pi}_{ji} \tag{A.23}$$

Take i, j. For $i = j$ the equation always holds so assume $i \neq j$. Then the condition reads

$$\begin{array}{ll} e^{-\beta E_i} P_{ij} e^{-\beta(E_j - E_i)} = e^{-\beta E_j} P_{ji} & for \quad \Delta E > 0 \\ e^{-\beta E_i} P_{ij} = e^{-\beta E_j} P_{ji} e^{-\beta(E_i - E_j)} & for \quad \Delta E \leq 0 \end{array} \tag{A.24}$$

All in all we have detailed balance if and only if $P_{ij} = P_{ji}$. We must in other words have a *symmetric* P. This means that if j is a neighbour of i ($P_{ij} \neq 0$) then i must be a neighbour of j ($P_{ji} \neq 0$). For the standard Metropolis algorithm a symmetric P means that two things are required: If $i \in N(j)$ then $j \in N(i)$ and the number $|N(j)|$ of neighbours of state j must be the same for all states j.

Normally we have the Gibbs distribution as the only stationary state because we normally have an *irreducible* process, which for discrete time means that the transition matrix has $\forall i, j : \tilde{\Pi}_{ij}^{(n)} > 0$ for some power n of $\tilde{\Pi}$, i.e. from any state it is possible to reach any other state through a series af transitions.

A.7 Glauber and others

I will now describe some generalizations and alternatives to the standard Metropolis algorithm. First, it is not necessary that all states have equal numbers of neighbours and to choose among them with equal probabilities. An example of a proposal distribution like that is shown in fig. A.1

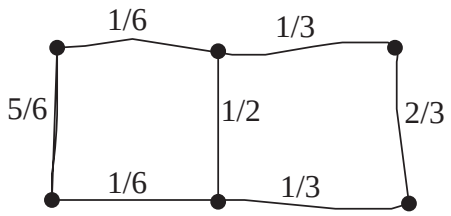


Figure A.1: A symmetric stochastic 6×6 matrix P shown as a graph with weighted bonds

Second, the acceptance probabilities of new states can be generalized. In general we must have

$$P_{accept}(\Delta E) = \begin{cases} f(\Delta E) & for \quad \Delta E \leq 0 \\ f(-\Delta E)e^{-\beta \Delta E} & for \quad \Delta E > 0 \end{cases} \tag{A.25}$$

f is some arbitrary function: $f :]\infty, 0] \rightarrow]0, 1]$. It is clear that the smaller values f has the smaller is the acceptance probability and it is more likely that it decides just to stay in its current state. The Metropolis algorithm with $f \equiv 1$ has therefore the fastest dynamics.

For a choice of f we could for example want a kind of antisymmetry:

$$\begin{array}{c} P_{accept}(\Delta E) = 1 - P_{accept}(-\Delta E) \\ \Updownarrow \\ f(-x)e^{-\beta x} = 1 - f(-x) \qquad for \quad x > 0 \\ f(x) = 1 - f(x)e^{\beta x} \qquad for \quad x < 0 \end{array} \tag{A.26}$$

i.e.

$$f(x) = \frac{1}{1+e^{\beta x}} \quad for \quad x < 0 \tag{A.27}$$

The resulting dynamics is known as *Glauber* dynamics. f has a nice property namely that

$$x > 0 : \qquad f(-x)e^{-\beta x} = \frac{e^{-\beta x}}{1+e^{-\beta x}} = \frac{1}{1+e^{\beta x}}$$

so we can simply write

$$P_{accept}(\Delta E) = \frac{1}{1+e^{\beta \Delta E}} \tag{A.28}$$

and

$$\tilde{\Pi}_{ij} = \begin{cases} \frac{P_{ij}}{1+e^{\beta \Delta E}} & for \quad j \neq i \\ 1 - \sum_{j \neq i} \tilde{\Pi}_{ij} & for \quad i = j \end{cases} \tag{A.29}$$

Sometimes we also rewrite the quantity

$$\begin{aligned} \frac{1}{1+e^{\beta \Delta E}} &= \frac{1}{2}\{1 + \tanh(-\frac{1}{2}\beta \Delta E)\} \\ &= \frac{1}{2}\{1 - \tanh(\frac{1}{2}\beta \Delta E)\} \end{aligned} \tag{A.30}$$

The Metropolis algorithm (or Glauber) is used in simulations because any initial distribution will tend to the Gibbs distribution. Therefore it is a way of sampling the Gibbs distribution which is useful in studies of physical systems in equilibrium. Because low energy states are most probable it is also used in optimization (simulated annealing). But it can also be used to study relaxation dynamics. The exact choice of f in eq. A.25 may be an attempt to model the dynamics of a real system as well as possible, or we may want to study a model dynamics in its own right.

A.8 Continuous time

Real systems have dynamics in continuous time, while the standard Metropolis algorithm (or whatever) works in discrete time. But it is easy to modify it to become a continuous time simulation. Instead of increasing the time t by 1 in each Metropolis step one can add

$$D = -\frac{1}{\lambda}\ln(x)$$

to time, where x is a random number uniformly distributed in the unit interval. The waiting time before a decision is taken will then have the following distribution

$$\begin{aligned} P(D > t) &= P(-\tfrac{1}{\lambda}\ln(x) > t) \\ &= P(\ln(x) < -\lambda t) \\ &= P(x < e^{-\lambda t}) \\ &= e^{-\lambda t} \end{aligned} \tag{A.31}$$

i.e. an exponential distribution with mean $1/\lambda$. This way one will make a Markov process in continuous time according to the scheme described in (A.2) where a decision may result in no jump. The rate of decisions λ'_i is the same and equal to λ for all states i. The "new" state is chosen according to the "jump" probabilities π'_{ij} where π'_{ij} is simply given by eq. A.22, so π' is simply the transition matrix $\tilde{\Pi}$ of the discrete time process.

But does this corresponding continuous time process still have the desired qualities? Yes: If the discrete time process is irreducible then the continuous time process is also irreducible because for all i: $\lambda_i = \lambda > 0$, and from any state i it is possible to reach any other state j via transitions through a series of states. Because if a discrete time transition $i \to k$ is possible then $\tilde{\Pi}_{ik} > 0$ and then $\pi'_{ik} > 0$ and then $\pi_{ik} > 0$ so it is also possible in continuous time. And the Gibbs distribution is still stationary, as we can see if we check detailed balance. Take $i \neq j$.

$$\begin{aligned} p_i^{Gibbs} W_{ij} &= p_j^{Gibbs} W_{ji} \\ \Leftrightarrow p_i^{Gibbs} \lambda_i \pi_{ij} &= p_j^{Gibbs} \lambda_j \pi_{ji} \\ \Leftrightarrow p_i^{Gibbs} \lambda'_i \pi'_{ij} &= p_j^{Gibbs} \lambda'_j \pi'_{ji} \\ \Leftrightarrow p_i^{Gibbs} \lambda \tilde{\Pi}_{ij} &= p_j^{Gibbs} \lambda \tilde{\Pi}_{ji} \\ \Leftrightarrow p_i^{Gibbs} \tilde{\Pi}_{ij} &= p_j^{Gibbs} \tilde{\Pi}_{ji} \end{aligned} \tag{A.32}$$

which is just the discrete time detailed balance.

So it is easy to make a continuous time simulation instead of a discrete time. But discrete time does not have to be a bad thing. On the small timescale there is a difference between the continuous and discrete simulations. But often you are only interested in what goes on on a larger timescale, and here you may not see the difference, because a sum of a stochastic variable n times will be relatively very close to n times its average (the law of big numbers). Adding $1/\lambda$ instead of $-\ln(x)/\lambda$ in each step is also computationally faster.

Appendix B

Ising systems

Consider an Ising model with Hamiltonian

$$H = -\frac{1}{2}\sum_{i,j} J_{ij} s_i s_j - h_0 \sum_i s_i \tag{B.1}$$

where $\sum_{i,j} = \sum_i \sum_{j \neq i}$ and the spins are up ($s_j = +1$) or down ($s_j = -1$). J_{ij} is the interaction strength (energy) between spins i and j and $J_{ij} = J_{ji}$. h_0 is the external field. The local (internal) effective field at site i is

$$h_i^{eff} = h_0 + \sum_{j \neq i} J_{ij} s_j \tag{B.2}$$

The magnetization is $m = \frac{1}{N}\sum_i s_i$ where N is the number of spins.

It is common in MC simulations to have single-spin-flips as allowed transitions. The energy increase in flipping spin i from s_i^{before} to s_i^{after} is

$$\begin{aligned}
\Delta E &= \Delta 2(-\frac{1}{2}\sum_{j \neq i} J_{ij} s_i s_j) - h_0 \Delta s_i \\
&= -\sum_{j \neq i} J_{ij} s_j \Delta s_i - h_0 \Delta s_i \\
&= \Delta s_i (-\sum_{j \neq i} J_{ij} s_j - h_0) \\
&= (s_i^{after} - s_i^{before})(-h_i^{eff}) \\
&= 2 s_i^{bef} h_i^{eff} = -2 s_i^{after} h_i^{eff}
\end{aligned} \tag{B.3}$$

The increase in magnetization is

$$\Delta m = \frac{2}{N} s_i^{after} = -\frac{2}{N} s_i^{before} \tag{B.4}$$

and

$$\Delta h_j^{eff} = J_{ij} \Delta s_i = 2 J_{ij} s_i^{after} = -2 J_{ij} s_i^{bef} \tag{B.5}$$

For Metropolis dynamics the acceptance probability is (see eq. A.20)

$$P_{accept} = \begin{cases} 1 & for \quad s_i^{bef} h_i^{eff} \leq 0 \\ e^{-2\beta s_i^{bef} h_i^{eff}} & for \quad s_i^{bef} h_i^{eff} > 0 \end{cases} \tag{B.6}$$

For Glauber dynamics the acceptance probability is (see eq. A.28)

$$P_{accept} = \frac{1}{1 + e^{2\beta s_i^{bef} h_i^{eff}}} = \frac{1}{2}\{1 - \tanh(\beta s_i^{bef} h_i^{eff})\} \tag{B.7}$$

For Glauber dynamics the probability that s_i is equal to ± 1 after it has been updated in the local field h_i is

$$\begin{aligned} \rho_{\pm 1} &= P(s_i^{after} = \pm 1) \\ &= P(s_i^{bef} = \mp 1) P_{accept}(\mp 2h_i) + P(s_i^{bef} \pm 1)(1 - P_{accept}(\pm 2h_i)) \\ &= P_{accept}(\mp 2h_i) = \frac{1}{1 + e^{\mp 2\beta h_i}} \end{aligned} \tag{B.8}$$

where we used eq. A.26.

B.1 h_{max}

Consider now ferromagnetic interactions $J_{ij} \geq 0$. At $T = 0$ it is possible for a spin i to point in the opposite direction than the external field h_0, if sufficiently many of its neighbors also are opposite h_0, i.e. if

$$s_i h_i^{eff} = s_i (h_0 + \sum_{j \neq i} J_{ij} s_j) > 0 \tag{B.9}$$

The largest possible magnitude of the ferromagnetic 'force' on spin i is $\sum_{j \neq i} J_{ij}$ so if $|h_0| > \sum_{j \neq i} J_{ij}$ then the external field is strongest and s_i will

have to align with h_0. I define h_{max} as the smallest magnitude of the external field h_0 where h_0 always will be stronger than internal interactions:

$$h_{max} = \max_i \sum_{j \neq i} J_{ij} \tag{B.10}$$

For $|h_0| \geq h_{max}$ the system is quite trivial since every spin which is opposite h_0 will decrease the energy when flipped, so at $T = 0$ the system soon will relax to the uniform ground state aligned to h_0. The largest magnitude of h_0 where the system is non trivial is therefore h_{max}.

Another interpretation is this: Since $\Delta E_i = 2 s_i^{bef} h_i^{eff}$ we see that the largest possible increase in energy associated with a spin flip when $h_0 = 0$ is $2h_{max}$. And this occurs when spin $k = \arg\max_i \sum_{j \neq i} J_{ij}$ is flipped opposite a uniform background of neighbors (a situation that is likely in nucleation). If all sites are equivalent in the model then $h_{max} = \sum_{j \neq i} J_{ij}$ for all i.

B.2 Weiss' mean field theory

It can be shown that in equilibrium (with the Gibbs distribution eq. A.18) the mean magnetization of spin i is

$$\langle s_i \rangle = \langle \tanh(\beta h_i) \rangle \tag{B.11}$$

In the *mean field approximation* we simply let s_i interact with the constant mean local field, in other words

$$\begin{aligned} \langle s_i \rangle &= \tanh(\beta \langle h_i \rangle) \\ &= \tanh[\beta(h_0 + \sum_{j \neq i} J_{ij} \langle s_j \rangle)] \end{aligned} \tag{B.12}$$

If all sites are equivalent (only ferromagnets) then $\langle s_i \rangle$ is independent of i and equal to the mean magnetization m, so

$$m = \tanh[\beta m (\sum_{j \neq i} J_{ij}) + \beta h_0] \tag{B.13}$$

where $\sum_{j \neq i} J_{ij}$ is independent of i. We have $h_{max} = \sum_{j \neq i} J_{ij}$. Possible states or phases have mean magnetizations which are solutions to eq. B.13. These are sketched for different temperatures in fig. B.1 at $h_0 = 0$. At the

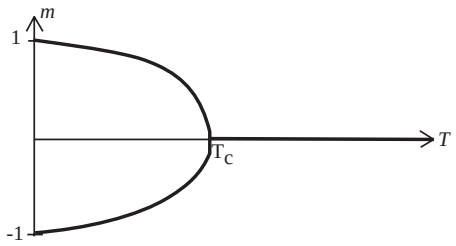


Figure B.1: Possible magnetizations at different temperatures.

critical temperature $T_c^{Weiss} = 1/\beta_c^{Weiss}$ and above, the thermal fluctuations win over the ferromagnetic ordering forces so there is only one solution $m = 0$. T_c^{Weiss} is given by $\beta_c^{Weiss}(\sum_{j\neq i} J_{ij}) = 1$ i.e. $T_c^{Weiss} = \sum_{j\neq i} J_{ij}$. We see that $T_c^{Weiss} = h_{max}$. Another thing that characterizes T_c^{Weiss} is that, if the interaction strength between two spins is taken to be of the order of one, then T_c^{Weiss} is the effective number of spins that one spin interacts with (the *coordination number*).

Bibliography

[Outbreak] Natalie S. Glance and Bernardo A. Huberman. The Outbreak of Cooperation. *Journal of Mathematical Sociology*, 17(4):281-302, 1993.

[Diversity] Bernardo A. Huberman and Natalie S. Glance. Diversity and collective action. In H. Haken and A. Mikhailov, editors, *Interdisciplinary Approaches to Complex Nonlinear Phenomena*. Springer, New York, 1993.

[Beliefs] Bernardo A. Huberman and Natalie S. Glance. Beliefs and Cooperation. In *Modelling Rational and Moral Agents*, P. Danielson ed., Oxford University Press, 1995.

[Behavior] Bernardo A. Huberman and Tad Hogg. The behavior of computational ecologies. In B. A. Huberman, editor, *The Ecology of Computation*, pages 77-115. North Holland, Amsterdam, 1988.

[Persistence] H. A. Ceccatto and B. A. Huberman. Persistence of nonoptimal strategies. *Proc. Natl.Acad. Sci.* USA, 86:3443-3446, 1989.

[Fluidity] Natalie S. Glance and Bernardo A. Huberman. Social dilemmas and fluid organizations. In K. Carley and M. Prietula, editors, *Computational Organization Theory*. Lawrence Elbraum Associates, Inc., New Jersey, 1994.

[SciAm] Natalie S. Glance and Bernardo A. Huberman. Dynamics of social dilemmas. *Scientific American*, March 1994.

[cellular] Bernardo A. Huberman and Natalie S. Glance. Evolutionary games and computer simulations. *Proc. Natl.Acad. Sci.* USA, 90:7716-7718, August 1993.

[Micro] S. H. Clearwater, T. Hogg, and B. A. Huberman. Cooperative problem solving. In *Computation: The micro and the macro view*, B. A. Huberman, editor, World Scientific, 1992, pp. 33-70.

[Training] Natalie S. Glance, Tad Hogg and Bernardo A. Huberman. Training and turnover in organizations. *Journal of Organization Science*, in press.

[Collaboration] Bernardo A. Huberman and Christoph H. Loch. Collaboration, motivation, and the size of organizations. Submitted to ?

[Multiagent] Natalie S. Glance and Tad Hogg. Dilemmas in Computational Societies. In V. Lesser, General Chair, *First International Conference on Multi-Agent Systems*, pages 117-124. AAAI Press/The MIT Press, Menlo Park, CA, 1995.

[Axelrod] Robert Axelrod and William D. Hamilton. The evolution of cooperation. *Science*, 211:1390-1396, 1981.

[MaynardSmith] John Maynard Smith. *Evolution and the theory of games*, Cambridge University Press, 1982.

[Lindgren] Kristian Lindgren. Evolutionary Phenomena in Simple Dynamics. In C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, editors, *Artificial Life II*, pages 295-312. Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. X. Addison-Wesley, Redwood City, CA, 1991.

[Lomborg] Bjørn Lomborg. *Simulating social science: The iterated prisoner's dilemma and computer simulations in political science.* Ph.D thesis, Institut for statskundskab, University of Copenhagen, licentiatsserien 1994/2.

[Altruism] J. R. Krebs and N. B. Davies. *An introduction to behavioural ecology*, 3.ed., chap.11, Blackwell Scientific Publications, 1993.

[Olson] Mancur Olson. *The Logic of Collective Action*. Harvard University Press, Cambridge, 1965.

[Bendor] Jonathan Bendor and Dilip Mookherjee. Institutional structure and the logic of ongoing collective action. *American Political Science Review*, 81(1):129-154, 1987.

[Ostrom] Elinor Ostrom. *Governing the commons: The evolution of institutions for collective action*, N.Y. Cambridge University Press, 1990.

[Network] Knipscheer and Antonucci, editors, *Social network research: Substantive issues and methodological questions*. Swets and Zeitlinger publishers, 1990.

[Burt] R. S. Burt. Models of network structure. *Annual review of Sociology*, Vol. 6, 79 (1980).

[Tragedy] Garrett Hardin. The tragedy of the commons. *Science*, 162:1243-1248, 1968.

[van Kampen] N. G. van Kampen. *Stochastic processes in physics and chemistry*, North Holland, Amsterdam, 1981.

[Markovkæder] Martin Jacobsen and Niels Keiding. *Markovkæder*, 3.ed., Institut for Matematisk Statistik, København, June 1990.

[Neural] John A. Hertz, Anders S. Krogh and Richard G. Palmer. *Introduction to the theory of neural computation*, Addison-Wesley, 1991.

[Glasses] K. H. Fischer and J. A. Hertz. *Spin glasses*, Cambridge University Press, 1991.

[Critical] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman. *The theory of critical phenomena*, Clarendon Press, Oxford, 1992.

[Annealing] Jacob Mørch Pedersen. *Simulated annealing and finite time thermodynamics.* Ph.D. thesis, Physics laboratory, H. C. Ørsted Institute, University of Copenhagen, 1990.

[Dyson] Freeman J. Dyson. Existense of a Phase-Transition in a One-Dimensional Ising Ferromagnet. *Commun. Math. Phys.* 21: 269-83, 1971.

[1] P. Sibani and J. A. Hertz. Parisi function for two spin glass models. *J. Phys. A: Math. Gen.*, 18: 1255-1274, 1985.

[2] G. J. Rodgers and A. J. Bray. Critical behaviour of Dyson's hierarchical model with a random field.

[3] George A. Baker, Jr. Ising model with a Scaling Interaction. *Phys. Rev.* B 5: 2622-2633, 1972.

[4] D. Kim and C. J. Thompson. Critical properties of Dyson's hierarchical model. *J. Phys. A: Math. Gen.* 10: 1579-1598, 1977.

[5] P. M. Bleher and Ya. G. Sinai. *Commun. Math. Phys.* 45: 247, 1975.

[6] Alba Theumann. Critical properties of the random hierarchical model. *Phys. Rev.* B 21: 2984-2990, 1980.

[7] D. L. Stein and C. M. Newman. A new look at broken ergodicity. September 1994.

[8] K. H. Hoffmann and P. Sibani. Diffusion in hierarchies. *Phys. Rev.* A 38: 4261-4270, 1988.

[9] Per Arne Rikvold and Bryan M. Gorman. Recent Results on the Decay of Metastable Phases. July 21, 1994.

[10] William Klein. *Critical phenomena and phase transitions.* Lecture notes, Ørsted laboratory, University of Copenhagen, 1992.

[11] A. B. Bortz, M. H. Kalos and J. L. Lebowitz. A New Algorithm for Monte Carlo Simulation of Ising Spin Systems. *J. of Comp. Phys.* 17: 10-18, 1975.

[12] M. A. Novotny. Monte Carlo algorithm with absorbing Markov chains: Fast local algorithms for slow dynamics. *Phys. Rev. Lett.* 74: pp. 1-5, 1995.