

Theory of Complex Systems

January 2019

Henrik Jeldtoft Jensen
Centre for Complexity Science and Department of Mathematics
Imperial College London

Contents

| | |
|---|-----------|
| 1 About this set of notes and recommended reading material | 4 |
| 2 Introduction | 6 |
| 3 Avalanche response in complex systems | 9 |
| 3.1 Sandpile models | 9 |
| 3.1.1 Nearest neighbour models | 9 |
| 3.1.2 Mean field models | 11 |
| 3.2 Forest fire | 13 |
| 3.2.1 Definition of the Forest Fire Model | 14 |
| 3.2.2 Simulation results for the Forest Fire model | 15 |
| 3.2.3 Mean Field discussion of the Forest fire model | 16 |
| 3.2.4 Physical Relevance of the Forest Fire Model | 19 |
| 4 Branching processes: generator formalism | 20 |
| 4.1 Generator Functions: sizes and lifetimes | 22 |
| 4.1.1 Size of the progeny | 26 |
| 4.1.2 Time to extinction | 29 |
| 5 Networks: generator formalism | 34 |
| 5.0.1 Cluster sizes | 36 |
| 5.0.2 The giant cluster | 39 |
| 6 Intermittency: Extreme Value Statistics, Record Dynamics and Tangent Map | 42 |
| 6.1 Extremal Value Statistics | 43 |
| 6.2 Records in White Noise | 49 |
| 6.3 Records and First Passage in Brownian Motion | 52 |
| 6.3.1 Relation between Record and Extreme Value Statistics | 54 |

| | | |
|-----------|--|------------|
| 6.4 | Summary | 56 |
| 6.5 | Tangent map intermittency | 57 |
| 6.6 | The Tangled Nature model | 58 |
| 6.7 | Analysis of the map in the neighbourhood of tangency | 65 |
| 7 | Percolation | 69 |
| 8 | Time correlations | 75 |
| 8.1 | Switching signal | 75 |
| 8.2 | 1/f signals | 77 |
| 8.2.1 | Transport by Diffusion | 78 |
| 8.3 | Non-stationary signals: Hurst exponent | 83 |
| 8.3.1 | Random walk with persistence or anti-persistence | 86 |
| 9 | Synchronisation | 92 |
| 10 | Statistical Mechanics: From individual component to systems level | 99 |
| 10.1 | How to obtain the probability measures | 100 |
| 10.2 | Ising model and mean field | 103 |
| 10.3 | The peculiar nature of a critical point | 106 |
| 11 | Emergence: From Micro to Macro - Vortices as a case study | 111 |
| 11.1 | The Two Dimensional XY-Model | 111 |
| 11.2 | Lack of Ordering in Two Dimensions | 116 |
| 11.3 | Vortex Unbinding | 117 |
| 11.3.1 | The Spin Wave Stiffness | 118 |
| 11.3.2 | The KT transition | 120 |
| 11.4 | The Vortex Unbinding Transition in Other Systems | 121 |
| 12 | Forecasting | 124 |

Chapter I

About this set of notes and recommended reading material

The present set of notes represent a personal view on what complexity science is about. They are intended as a companion to the lectures *The Theory of Complex Systems* delivered during the spring term as part of the master courses offered by the Department of Mathematics at Imperial College London.

The material below draws on previous lecture notes and published papers and the following two books:

- *Self-organized Criticality. Emergent Complex Behavior in Physical and Biological Systems*, Henrik Jeldtoft Jensen, Cambridge University Press, 1998.
- *Stochastic Dynamics of Complex Systems. From Glasses to Evolution*, Paolo Sibani and Henrik Jeldtoft Jensen, Imperial College Press, 2013.
- How Statistical Mechanics Describes Emergent Phenomena, Henrik Jeldtoft Jensen in Meyers: Encyclopedia of Complexity and Systems Science. Ed. R.A. Myers, 2009.

There are very many relevant books one may benefit from consulting. A super short, though still meaningful, account is given by John H Holland in his *COMPLEXITY. A Very Short Introduction*, Oxford. Below I simply list a few treatises I personally find very informative:

- Gunnar Pruessner: Self-Organised Criticality, Theory, Models and Characterisation. Cambridge, 2012.

- Didier Sornette: Critical Phenomena in Natural Sciences, Chaos, Fractals, Self-organization and Disorder: Concepts and Tools. Springer, 2004.
- Theodore E Harris: The Theory of Branching Processes, Dover, 1989.
- Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths: Synchronization, A universal concept in nonlinear sciences. Cambridge, 2001.
- Edward Ott: Chaos In Dynamical Systems, Cambridge 2002.
- Steven H Strogatz: Nonlinear Dynamics and Chaos. Westview, 1994.

and as supplement:

- Manfred Eigen, From Strange Simplicity to Complex Familiarity, A Treatise on Matter, Information, Life and Thought. Oxford, 2013.
- Neil Johnson, Simply Complexity, Oneworld Publications, 2009.
- Melanie Mitchell, Complexity: A Guided Tour, Oxford 2009.
- M Mitchell Waldrop, Complexity: The Emerging Science at the Edge of Order and Chaos, Penguin, 1994.
- John J. Holland: Complexity: A Very Short Introduction. Oxford University Press, 2014.

Chapter 2

Introduction

Theory of complexity is about how at the level of systems interconnectedness and (hierarchical) structure emergence.

Our focus will be on the behaviour at the systemic level and how it differs from the level of the individual components, often referred to as agents. To make clear the relationship between the dynamics at the level of the individuals and the collective systemic level let us consider two examples. First think of the ideal gas. It may consist of about $N = 10^{23}$ particles, but since the particles do not interact the statistics of a single particle is identical to the statistics of the entire gas in the sense that pressure and temperature are given directly in terms of single particle averages. The temperature, as is well known, is equal to the average kinetic energy of the individual molecules according to $\frac{3}{2}k_B T = \langle \frac{1}{2}mv^2 \rangle$ and the pressure is N times the average momentum transfer per area of each particle. Since the particles do not interact, the collective systemic behaviour is obtained by summing up the contributions from each component. No new emergent phenomena arise.

The situation is different when components can interact. The Schelling model[64] is a particular telling example of the difference that can occur between the behaviour of the individual and the collective. Schelling was interested in segregation in cities and noticed that individual citizens typically is quite tolerant and can happily accept to live in a suburb where their ethnicity constitutes a minority. Say black people when asked individually are willing to stay in a neighbourhood even 75% of the local population is white. Nevertheless, segregation into pure white and pure black neighbourhoods are frequently observed rather than some mix of $X\%$ white and $(100 - X)\%$ black. The question is why the expectation one would get by extrapolating the tolerance of the individuals does not predict the observations at the level of the community.

Schelling's considers agents of two types A and B moving around on a lattice. The

agents have a preference to reside in a neighbourhood inhabited by agents of the same type as themselves, i.e. A prefers A and B prefers B . But the agents are also tolerant and willing to stay in a neighbourhood where the other type constitutes a certain percentage of the population. If the tolerance of the individual controlled the collective dynamics, the model should exhibit regions of a mixture of A and B agents residing together. Instead, one observes at the macroscopic level segregation of the population into purely A or purely B regions, even if the individual agent possesses a tolerance level that makes a agent willing to remain in a region as long as like agents constitute 33% or more of the local population. This finding is a dramatic example of how emergent systemic behaviour can be very different from the behaviour of the individual components. We shall return to this theme through out the course.

But first let us note that the world consists of inter-connected processes. It is an illusion to think that “things” exist. As argued by Alfred North Whitehead processes are ontologically fundamental. Since the building blocks do not consist of “things” with specific intrinsic properties, we can hope that a general science of the processes underlying and controlling the behaviour of apparently very different situations such as the evolution of an ecosystem or the performance of a piece of music by a band, may very well exist.

A saxophone and a tree doesn’t have much in common. But a jazz band and a forest might very well have. Say for instance in the way information is moving around amongst the components and in how structures evolve. The similarity may extend further. For instance, a saxophone is of course not really a “thing”. It is a collection of processes that move objects around and generates the process of variation in air pressure; which in our mind becomes the wonderful phenomena we call music. Similar for ecosystems. What we see as “things”, building blocks, or components at one level, are at another level themselves collections of components participating in processes.

Why is this, rather self-evident observation, important to us here. Because, as soon as one realises that the world is made of inter-connected processes and not of “things”; one also immediately realises why Complexity Science is the most fundamental of the sciences and why this scientific activity may very likely to create more fundamental insights about the world we are surround by than, say, the Large Hadron Collider ever will be able to.

How should complexity science go about uncovering the generalities behind the super complex systems such as the brain, the economy or ecologies? By finding ways to extend the tremendously successes of statistical mechanics, which is a very successful example of the science of multi-component systems and was developed around 1900 by people like Gibbs, Boltzmann and Einstein. At present, statistical mechanics has its greatest successes when applied to equilibrium systems. Complexity Science

needs to discover ways to systematically extend statistical descriptions to far from equilibrium. In doing so we must not be naïve, we must not be too attached to our old formalisms. But there are indications that ordinary statistical mechanics of systems at or near equilibrium may point us in the right direction. We need to expand our understanding of intermittency, tipping points, stability – in short: we need to understand the emergence of hierarchies. We need to identify which aspects of complex systems are typical across many types of systems and we need to learn what kind of mathematical formalism is able to capture and describe the emergent hierarchical structures.

Big Data will undoubtedly be a big help. But only if we use the access to big data to identify simple unifying concepts. We won't gain much insight by making our models as complicated and involved as the phenomena we try to understand or as data rich as our Big Data banks. We need understanding of co-evolution, simple networks, statistical mechanics and information theory.

Chapter 3

Avalanche response in complex systems

Since the introduction of the concept of Self-Organised Criticality (SOC) [53] by focusing on the dynamics of avalanches burst activity in a large number of systems have been investigated with the anticipation that the distribution of sizes of the bursts can be used as an indicator of whether the system is critical (in the sense of diverging correlations lengths and time) or not. For a brief introduction to SOC see [35] and for a recent very thorough discussion of SOC see the wonderful book by Gunnar Pruessner [57]. We'll just look at how branching processes may be involved in generating, at least approximatively, the often observed distribution functions.

3.1 Sandpile models

3.1.1 Nearest neighbour models

We briefly present the so called d -dimensional sandpile model first introduced by Bak, Tang and Wiesenfeld in 1987 when they introduced the concept of Self-Organised Criticality (for more detail see [35] and [57]). We define a dynamical variable $z(\mathbf{r})$ on a cubic d -dimensional lattice of linear size L . The basis vectors of the lattice is denoted by \mathbf{e}_i with $i = 1, 2, \dots, d$.

Operation rules:

R1 *Conservative perturbation*

$$\begin{aligned} z(\mathbf{r}) &\rightarrow z(\mathbf{r}) + d \\ z(\mathbf{r} - \mathbf{e}_i) &\rightarrow z(\mathbf{r} - \mathbf{e}_i) - 1 \text{ for } i = 1, \dots, d \end{aligned} \tag{3.1}$$

R2 *Nonconservative perturbation*

$$z(\mathbf{r}) \rightarrow z(\mathbf{r}) + 1 \tag{3.2}$$

R3 *Relaxation*

If $z(\mathbf{r}) > z_c$, then

$$\begin{aligned} z(\mathbf{r}) &\rightarrow z(\mathbf{r}) - 2d \\ z(\mathbf{r} \pm \mathbf{e}_i) &\rightarrow z(\mathbf{r} \pm \mathbf{e}_i) + 1 \text{ for } i = 1, \dots, d \end{aligned} \tag{3.3}$$

R4 *Boundary conditions*

We visualise sand piling up against a corner. The sand is able to drop off over the edges opposite to the corner of the cube.

Closed boundary – The dynamical variable is kept equal to zero along the one corner of the cube, that is for positions $\mathbf{r} = (r_1, r_2, \dots, r_d)$ for which at least one of the coordinates r_i is equal to zero. I.e., we have

$$z(\mathbf{r}) = 0 \text{ if there exist a co-ordinate } r_i = 0 \tag{3.4}$$

Open boundary – Sand leaves over the edges defined by having one of the co-ordinate r_i equal to L . If $z(\mathbf{r}) > z_c$ for \mathbf{r} along this corner we perform the following update:

$$\begin{aligned} z(\mathbf{r}) &\rightarrow z(\mathbf{r}) - 2d + \text{number of } i \text{ with } r_i = L \\ z(\mathbf{r} + \mathbf{e}_i) &\rightarrow z(\mathbf{r} + \mathbf{e}_i) + 1 \text{ if } r_i \neq L \\ z(\mathbf{r} - \mathbf{e}_i) &\rightarrow z(\mathbf{r} - \mathbf{e}_i) + 1 \text{ for } i = 1, \dots, d. \end{aligned} \tag{3.5}$$

The lattice is updated simultaneously. If more than one site become over-critical, $z(\mathbf{r}) > z_c$, we update all these sites at the same time. A time step is defined as a visit to all the L^d sites checking whether their z -value is larger than z_c or not and performing an update whenever $z(\mathbf{r}) > z_c$. The pile is perturbed by rule R1 or R2. If over-critical sites are produced then relaxation is performed according to rule R3 and R4 until all sites again are subcritical.

3.1.2 Mean field models

To present a mean field discussion of avalanche dynamics in models of complex systems such as the sand pile, forest fires and more recent models of fusions of banks.

Consider a random neighbour version of the BTW sandpile cellular automaton described in detail in [35] or [57]. The effect of a random assignment of neighbours instead of the usual nearest neighbour arrangement is to destroy spatial correlations. We will in this sense consider the random neighbour model as a kind of mean field theory.

The model is defined as in Sec. 3.1.1 except that we choose $q_c = \alpha z_c$ new neighbours at random every time a site topples (Christensen and Olami 1993). Here q_c is the coordination number of the site and α is the conservation level of the model, this we explain now. Each of the randomly chosen sites receive one unit of sand $z_{rn} \rightarrow z_{rn} + 1$. The overcritical site, say site number i , loses z_c grains of sand, i.e., $z_i \rightarrow z_i - z_c$. Notice that an element of dissipation has been introduced. Since only αz_c sites receive a grain of sand the model is non-conservative except in the limit $\alpha \rightarrow 1$. We imagine that the model is driven by adding extra units of sand at random positions (This is rule R2 above).

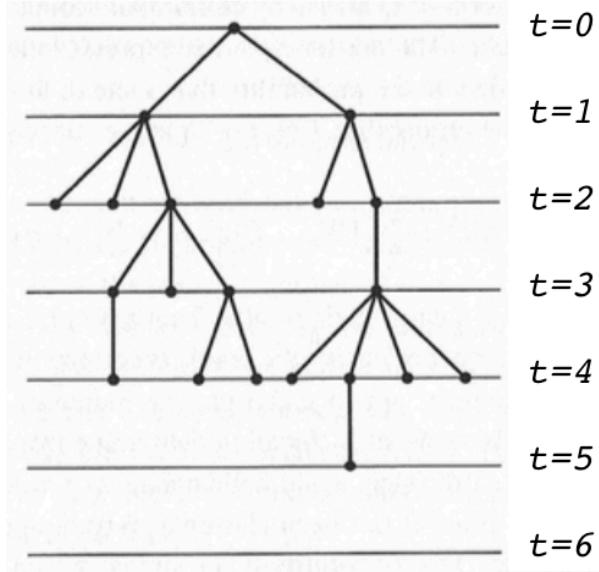


Figure 3.1: Branching process

It is very illuminating to think of the evolution of the avalanches in the model as a

kind of branching process [31]. We'll discuss branching processes in more mathematical detail in Sec. 4 Fig. 3.1 illustrates the tree like structure created by a branching process. The process starts at the top. Each node branch off a number n of branches. The probability that a node divides out into $n = 0, 1, 2, \dots$ branches is denoted by p_n . A toppling site can induce from zero to q_c new topplings. The temporal evolution of an avalanche can be considered as a branching process in which from zero to q_c new branches can emerge. The average number of new branches is called the branching ratio $\sigma = \langle p_n \rangle$. The probability that a tree in a branching process of branching ratio σ develops s nodes before it dies is given by (Harris 1963)

$$P(s) \propto s^{-3/2} \exp\left(-\frac{s}{s_0(\sigma)}\right). \quad (3.6)$$

The exponential cut-off diverges when $\sigma \rightarrow 1^-$ according to $s_0(\sigma) \propto (1 - \sigma)^{-2}$. The tree size distribution is equal to the avalanche size distribution when we consider the sandpile as a branching process.

The branching probability is easily expressed in terms of the probability density $P(z)$. Here $P(z)$ denotes the probability that the dynamical variable z_i of an arbitrarily chosen site i assumes the value z , where z is an integer. A site becomes over-critical when $z_i \geq z_c$. Hence, the q_c grains distributed when a site topples will induce new topplings (or new branches) when they hit sites with z -values equal to $z_c - 1$. The probability that a toppling induces n new topplings are therefore given by [17]

$$p_n = K_{q_c, n} P^n(z_c - 1) (1 - P(z_c - 1))^{q_c - 1}. \quad (3.7)$$

Here $K_{i,j}$ denotes the binomial coefficient. The average branching ratio σ is given by

$$\sigma = \langle n \rangle = \sum_{n=0}^{q_c} np_n = q_c P(z_c - 1). \quad (3.8)$$

The probabilities $P(z)$ for $z < z_c$ are easily determined from their master equation in the stationary state. The master equation expresses the change of $P(z)$ with time. Let $P(z, t)$ denote the probability that a site in the system assumes the value z at time t . The probability $P(z, t + 1)$ in the succeeding time step is given by

$$P(z, t + 1) = P(z, t) + \sum_{z'} [W_{z, z'; t} P(z', t) - W_{z', z; t} P(z, t)]. \quad (3.9)$$

The fraction of states that change their z -value from z' to z during a time step is denoted by $W_{z, z'; t}$. The first term in the bracket corresponds to sites changing their

z -value from some $z' \neq z$ to z . This process increases $P(z, t + 1)$. The second term corresponds to the sites changing their value from z to some other z -value. During a single time step a site will change its z -value by an amount equal to the number of times it is chosen as a (random) neighbour for a relaxing over-critical site. The probability that a site is chosen as neighbour is equal to $1/N$ in a system of size N . I.e. the probability that a site is chosen to be neighbour to k over-critical sites is equal to N^{-k} . In the limit of $N \rightarrow \infty$ we can neglect the possibility $k > 1$. In this approximation a site can at most receive one unit of sand during a time step. The transition probabilities connect z -values that differ precisely by one unit. When a site topples it induces a change of the z -value of q_c other sites, accordingly we have

$$W_{z,z';t} = \frac{q_c n_t}{N} \delta_{z-1,z'}, \quad (3.10)$$

where n_t is the number of topplings in the system at time t and $\delta_{i,j}$ denotes the Kronecker δ -function ($\delta_{i,j} = 0$ when $i \neq j$ and $\delta_{i,j} = 1$ when $i = j$). From equation 3.9 we obtain that

$$P(z, t + 1) - P(z, t) \propto [P(z - 1, t) - P(z, t)]. \quad (3.11)$$

Thus in the time independent stationary state where $P(z, t + 1) = P(z, t) = P(z)$ we conclude that $P(z - 1) = P(z)$ for all $z = 0, 1, 2, \dots, z_c - 1$. We normalise

$$\sum_{z=0}^{z_c-1} P(z) = 1 \quad (3.12)$$

and obtain $P(z) = 1/z_c$ for all $z < z_c$.

From Eq. 3.8 we obtain the branching ratio $\sigma = \alpha$. From the expression in Eq. 3.6 for the distribution of tree sizes we see that the non-conservative, $\alpha < 1$, random neighbour BTW model is non-critical. For avalanche sizes $s < s_0$ the avalanche size distribution behaves like a power law, but the exponential cut-off at $s_0(\alpha)$ is *independent* of N . This conclusion agrees with results from simulations of the random neighbour BTW model [17]. In the next section we discuss how the OFC model can remain critical even in the non-conservative case.

3.2 Forest fire

In this section we describe a model developed by Drossel and Schwabl [19]. The model is known as the forest fire model. The degree of relationship to real forest fires remains unknown. The model has, however, with some success been used to model the spreading of measles in the population of the island of Bornholm and of the Faroe Islands

[60]. To be consistent with the literature we will use the terminology of trees and fires. The harm done in doing so is no greater than in using sandpile language for the Bak, Tang and Wiesenfeld cellular automata (See Sec. 3.1.1).

We include the model in our discussion for several reasons. Firstly, the model is simple and pretty and may indeed throw some light on the burst like temporal fluctuations observed in epidemiology. Secondly, the numerical simulations indicate that the model, in a certain limit, drives itself to a state close to criticality[58]. Thirdly, a neat mean field theory exists for the model [16], which we'll discuss below.

3.2.1 Definition of the Forest Fire Model

The model is defined on a d -dimensional cubic lattice. Each site of the lattice can be in one of three different states.

- 1) The site can be empty.
- 2) The site can contain a green tree.
- 3) The site can contain a burning tree.

The lattice is updated in parallel according to the following algorithm:

- 1) A site occupied by a burning tree becomes an empty site in the succeeding time step.
- 2) A green tree becomes a burning tree if one or more of its nearest neighbour sites contain a burning tree.
- 3) An empty site becomes occupied by a green tree with probability p (the growth rate) in each time step.
- 4) A green tree which is not neighbour to burning sites catch spontaneously fire with probability f in each time step (the lightning rate).

Periodic boundary conditions are assumed and the initial configuration can, for instance, be a random configuration of green trees and empty sites.

Rule number 4 turns out to be crucial for the critical behaviour. If the lattice is updated according to the first three rules only, the evolution becomes periodic rather than critical. System spanning spiral-like fire fronts traverse the system with a period proportional to $1/p$. A characteristic scale develops when the spontaneous ignition is left out. This can be understood in the following way. Small clusters of green trees will

inevitably become large clusters since the only mechanism that can limit their size is that they grow until they hit and become connected to a cluster of trees which contains a burning front. A characteristic size of stable clusters of trees can be estimated by balancing the growth against the number of burning trees. For a cluster size to be stable the number N_b of trees burning away in each time step must be balanced by the number of new green trees N_g grown along the perimeter of the cluster during a time step. New trees grow along the empty sites along the perimeter of the considered cluster. Let N_p denote the number of sites along the perimeter. We then have $N_g \propto pN_p$. Since sites neighbour to burning sites catch fire with probability one the number of burning sites N_b will in each time step be of order one or larger. Thus, balance between N_b and N_g implies that N_p must increase like $1/p$ when $p \rightarrow 0$. In this way we see that a characteristic scale given by $1/p$ exist.

Numerical simulations as well as analytical analysis show that rule 4) in the updating algorithm secures the apparent critical behaviour in the double limit $p \rightarrow 0$ and $f/p \rightarrow 0$. The condition that $f \ll p$ is important. If the ignition probability is too high large-scale structures are never formed. Clusters of green trees simply burn down before they manage to grow large.

3.2.2 Simulation results for the Forest Fire model

All though careful large scale simulations makes it clear that the Forest Fire Model isn't critical in the ordinary sense of exhibiting finite size scaling that extrapolates to pure power law behaviour in the limit of infinite system size, the model is never the less sufficiently close to criticality that numerical simulations, at least approximately, exhibit many features of critical scale free behaviour.

Simulations of the model in 1 to 8 dimensions confirm the critical behaviour of the model [15]. The probability $P(s)$ that a cluster of green trees contains s trees exhibit power law behaviour $P(s) \sim s^{-\tau}$. Where $\tau = 1$ in one dimension and increases to $\tau = 1.5$ in six dimensions. The clusters of green trees are fractal objects in dimensions larger than 2. This is seen from the scaling of the radius of gyration $R(s)$. The radius of gyration for a cluster of size s is defined as the mean distance of the trees from the centre of mass of the cluster. That is

$$R(s) = \frac{1}{s} \sum_{i=1}^s |\mathbf{r}_i - \langle \mathbf{r}_i \rangle| \quad (3.13)$$

The simulations show that $R(s) \sim s^{1/\mu}$. This scaling indicates a fractal structure when $\mu < d$. The numerics show this to be the case for $d > 2$ and perhaps also for $d = 2$ [26, 15]. The simulations find that the scaling exponents change with dimension

for $d < 6$ for then to remain independent of dimension for $d \geq 6$. This suggests that the upper critical dimension of the model is identical to 6. This is in agreement with analytic mean field arguments involving a mapping of the forest fire model onto percolation, see Sec. next section.

3.2.3 Mean Field discussion of the Forest fire model

A mean field description of the forest fire model defined in Sec. 3.2 has been constructed by Christensen, Flyvbjerg, and Olami 1993 [16]. The theory consist of a set of equations of motion for the three densities characterising the state of the model. That is the density of trees $\rho_t(t)$, the density of trees in fire $\rho_f(t)$, and finally the density of empty sites $\rho_e(t)$. Assume that there are many more green trees than there are trees in fire. In this limit a tree can catch fire by interacting with a single burning site or by being hit by a lightning. The latter happens with rate f . The time evolution of the three densities is given by the following set of equations

$$\rho_e(t+1) - \rho_e(t) = -p\rho_e(t) + \rho_f(t) \quad (3.14)$$

$$\rho_t(t+1) - \rho_t(t) = -(f + q_c\rho_f(t))\rho_t(t) + p\rho_e(t) \quad (3.15)$$

$$\rho_f(t+1) - \rho_f(t) = -\rho_f(t) + (f + q_c\rho_f(t))\rho_t(t). \quad (3.16)$$

The change in the number of empty site (Eq. 3.14) consists of two terms. The first describes the loss of empty sites by the growing of new green trees at a rate p times the number of available empty sites. The second contribution consists of an addition of new empty sites created by the burning down of trees. The change in the number of trees (Eq. 3.15) consists of a depletion of tree sites by the spontaneous lightning and the possible spreading of fires from one¹ of the $q_c\rho_f$ neighbour sites in fire. New trees are added at the rate p times the number of available empty sites. Finally the change in the number of sites in fire, Eq. 3.16. The sites in fire at time t has become empty sites at time $t+1$. That explains the first term on the right hand side of Eq. 3.15. Green trees adds to the number of sites in fire at a rate f times the number of green trees able to be hit by a lightning. This is the second term of the equation. Finally the term entering Eq. 3.15 as a depletion term, namely the number of green trees catching fire from a burning neighbour sites, enters in Eq. 3.16 as a positive contribution. Since the sites of the model can be divided in to the three complementary categories *tree*, *fire*, or *empty* the densities must add up to one. I.e.,

$$\rho_t(t) + \rho_f(t) + \rho_e(t) = 1 \quad (3.17)$$

¹Remember we consider the limit of very few fires compared to trees

We can use this constraint to eliminate one of the variables in the dynamical equations (Eq. 3.14–3.16). Furthermore, we will consider time as a continuous variable and replace the time differences by derivatives. We then arrive at the following set of dynamical equations

$$\begin{aligned}\dot{\rho}_t &= -f\rho_t - q_c\rho_t\rho_f + p(1 - \rho_t - \rho_f) \equiv F(\rho_t, \rho_f) \\ \dot{\rho}_f &= -\rho_f + f\rho_t + q_c\rho_t\rho_f \equiv G(\rho_t, \rho_f).\end{aligned}\quad (3.18)$$

We can apply standard linear stability analysis to this set of equations. First we identify the fixed point of the equation by solving the equations $\dot{\rho}_t = 0$ and $\dot{\rho}_f = 0$. There is only one solution compatible with the requirement that $\rho_t \leq 1$. Namely,

$$\begin{aligned}\rho_t^* &= \frac{1}{2q_c}(1 + q_c + \kappa - [\kappa^2 + 2(1 + q_c)\kappa + (q_c - 1)^2]^{1/2}) \\ \rho_f^* &= \frac{p}{1 + p}(1 - \rho_t),\end{aligned}\quad (3.19)$$

where $\kappa \equiv f(1 + p)/p$. The behaviour in the vicinity of the fixed point (ρ_t^*, ρ_f^*) is determined by the eigenvalues, λ , of the determinantal equation

$$\begin{vmatrix} \partial F / \partial \rho_t - \lambda & \partial F / \partial \rho_f \\ \partial G / \partial \rho_t & \partial G / \partial \rho_f - \lambda \end{vmatrix} = 0. \quad (3.20)$$

This equation is of the form $\lambda^2 + B\lambda + C = 0$ with

$$\begin{aligned}B &= 1 + f + p + q_c(\rho_f^* - \rho_t^*) \\ C &= f(1 + p) + p(1 + q_c(\rho_f^* - \rho_t^*)) + q_c\rho_f^*.\end{aligned}$$

In the limit $f/p \rightarrow 0$ and $p \rightarrow 0$ the two roots of this parabola are complex with negative real parts. We conclude that the fixed point of Eqs. 3.2 is an attractive spiral. The phase plane trajectories are sketched in Fig. 3.2. The density of trees in the asymptotic regime is predicted to be equal to $\rho_t^* = 1/q_c$ in the limit $f = 0$.

The forest fire can also be considered as a branching process in the same way as the sandpile model was in Sec. 3.1.2. The branching ratio σ is equal to the average number of trees ignited by a fire spreading from one burning tree. A tree can ignite from 0 up to q_c of its neighbours. The calculation of σ is given by same expression as in Eq. 3.7 and Eq. 3.8. That is

$$\begin{aligned}\mu &= \sum_{n=0}^{q_c} n K_{q_c, n} \rho_t^n (1 - \rho_t)^{q_c - n} \\ &= 1 - \frac{\kappa}{q_c - 1} + O(\kappa^2).\end{aligned}\quad (3.21)$$

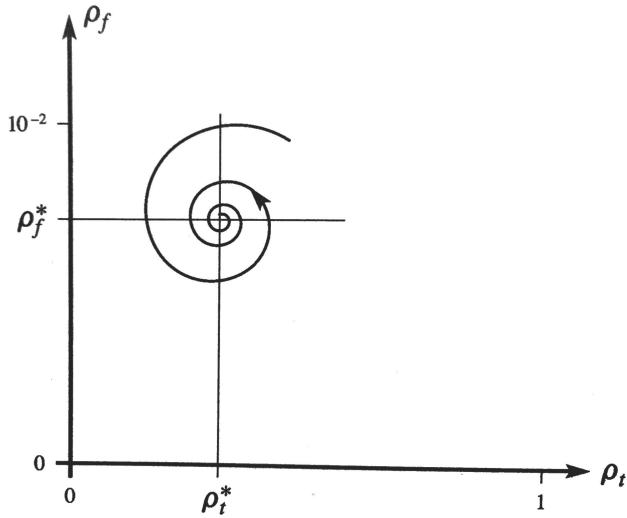


Figure 3.2: The phase plane trajectories of the mean field forest fire model.

I.e. the forest fire becomes a critical branching process in the limit $f \rightarrow 0$. The distribution of sizes of fires is given by Eq. 3.6. I.e. the value for the exponent of the size distribution is $\tau = 3/2$.

The forest fire can also be considered as a percolation process[16]. This description is still “mean field” in the sense that spatial correlations are neglected and trees are considered as uniformly randomly distributed over the lattice at a density equal to ρ_t . In this approximation the sizes of fires are simply given by the expression in percolation theory for the sizes of connected clusters [71]. What we mean is the following. Consider a lattice in d dimensions. Choose a site at random and plant a tree at the site. Continue this process until the fraction ρ_t of all sites contain a tree. A cluster is defined as a set of trees all connected through nearest neighbour links. If one tree in the cluster catches fire the whole cluster will eventually burn down since the fire will be able to spread through nearest neighbour links to all trees in the cluster. This way of considering the forest fire model allows one to make use of several results in percolation theory. In particular the result that the upper critical dimension is 6. This means that the values for power-law exponents derived in mean field theory should be correct in dimensions larger than or equal to 6. In fact simulations in 1 to 6 [16] and in 1 to 8 dimensions [15] confirm this expectation.

3.2.4 Physical Relevance of the Forest Fire Model

The dynamics of the Forest Fire Model has been related to the fluctuations in the temporal variation in the recorded number of measles cases is very accurately registered at the Danish island Bornholm as well as at the Faroe Islands. The number of cases fluctuates greatly. Bursts of varying sizes in the number of infected is followed by longer or shorter periods of quiescence. These islands are rather isolated and the development of an epidemic burst can perhaps be thought of as being driven by more or less internal interaction amongst the members of the population. New outbursts are probably to a large extend induced at random by visits from the outside. This is probably why the dynamics of the spreading of measles resembles the dynamics of the forest fire model. See [35] for details and references. More recently the dynamical activity of the brain has been discussed in terms of Forest Fire Model dynamics by Dante Chialvo and his collaborators. See e.g. Fig. 3 in [72].

Chapter 4

Branching processes: generator formalism

In this chapter we will discuss the mathematics of branching processes in some more detail making use of the generator function formalism. Let p_k denote the probability that a node has k out going branches.

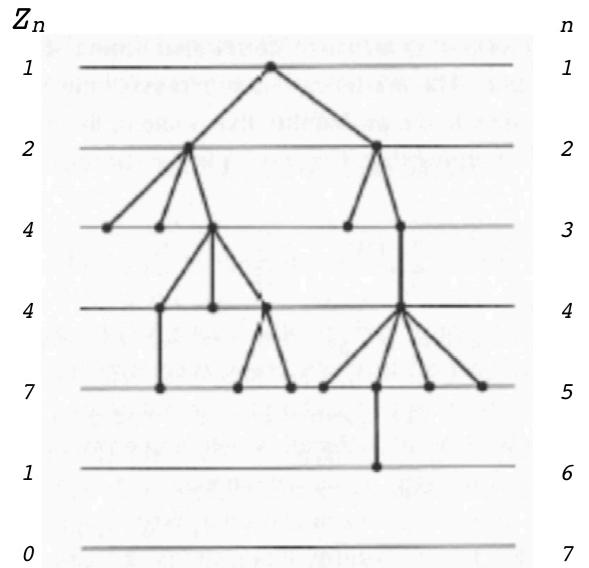


Figure 4.1: Branching process. The node in generation 0 has two branches. In generation 1 we see a node with 3 and one with 2 branches etc.

Normalisation demands

$$\sum_{k=0}^{\infty} p_k = 1 \quad (4.1)$$

and the average number of branches is given by

$$\mu = \sum_{k=0}^{\infty} kp_k. \quad (4.2)$$

Let Z_n denote the size of the population at generation n . We want to discuss the

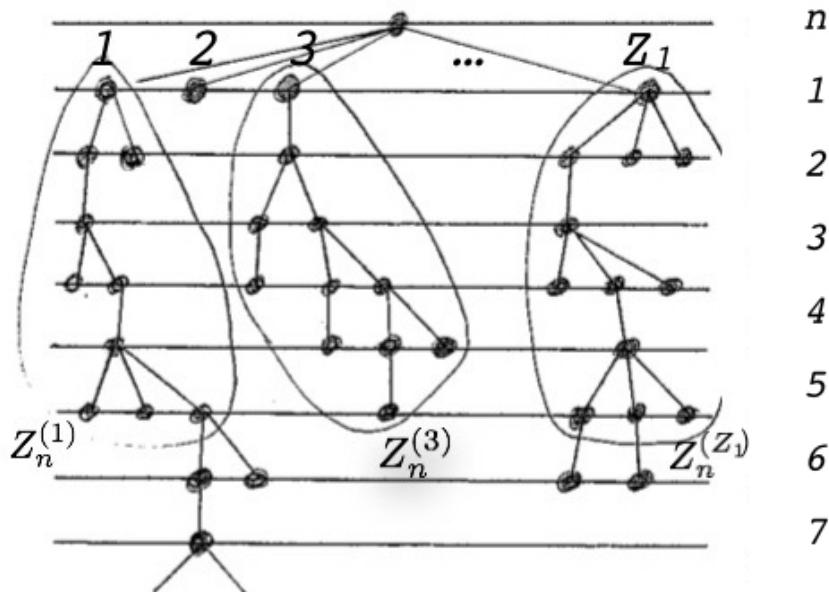


Figure 4.2: Branching process. The tree is broken up into Z_1 trees according to the prodigy of the first generation. Each of these sub-tress follows the same statistics as the entire tree.

following properties.

- (1) Extinction, i.e. $\text{prob}\{Z_n = 0\}$ for $n < \infty$.
- (2) The time dependence of the average size of the n -th generation $\langle Z_n \rangle$.
- (3) The distribution of tree sizes.

(4) The distribution of the time to extinction.

The self-similarity of the generated tree structures is the key property that allow one to make this analysis. At generation n we group the nodes into groups according to their parent at generation 1:

$$Z_n = \sum_{k=1}^{Z_1} Z_n^{(k)}. \quad (4.3)$$

Here $Z_n^{(k)}$ denotes the number of nodes in generation n that originates from node k in the first generation. There are Z_1 nodes in generation one and one node, per definition, in generation zero.

We'll analyse the statistics of these trees by use of *Generator functions*. In the next section we briefly introduce this formalism.

4.1 Generator Functions: sizes and lifetimes

Let $X = 0, 1, 2, \dots$ be a discrete stochastic variable and let

$$P_X(n) := \text{prob}\{X = n\}. \quad (4.4)$$

The generator function for X is defined as

$$g_X(s) = \sum_{n=0}^{\infty} P_X(n)s^n. \quad (4.5)$$

It is straight forward to check that we have

$$\begin{aligned} g_X(1) &= \sum_{n=0}^{\infty} P_X(n) = 1, \\ g'_X(1) &= \sum_{n=0}^{\infty} nP_X(n) = \langle X \rangle, \\ g''_X(1) &= \langle X^2 \rangle - \langle X \rangle. \end{aligned} \quad (4.6)$$

We often have to deal with the following type of sum

$$S_N = \sum_{k=1}^N X_k, \quad (4.7)$$

where X_k are *iid*, and N is a stochastic variable able to assume values in $\{0, 1, 2, \dots\}$. Let the generator for N be

$$g_N(s) = \sum_{n=0}^{\infty} P_n(n)s^n, \quad (4.8)$$

and the generator for all the X_k be the $g_X(S)$ above. In this case the following holds for the generator function for the sum S_N (see e.g. [21]):

$$g_{S_N}(s) = g_N(g_X(s)). \quad (4.9)$$

The proof goes as follows:

Proof

We have $g_{S_N}(s) = \sum_{q=0}^{\infty} P_{S_N}(q)s^q$ and

$$P_{S_N}(q) = \text{prob}\{S_N = \sum_{k=1}^N X_k = q\} \quad (4.10)$$

$$= \sum_{n=0}^{\infty} \text{prob}\{N = n\} \text{prob}\{\sum_{k=1}^n X_k = q\} \quad (4.11)$$

$$= \sum_{n=0}^{\infty} P_N(n) \sum_{q_1} \cdots \sum_{q_n} \delta_{q_1+\cdots+q_n, q} P_X(q_1) \cdots P_X(q_n) \quad (4.12)$$

and therefore

$$g_{S_N}(s) = \sum_{q=0}^{\infty} \sum_{n=0}^{\infty} P_N(n) \sum_{q_1} \cdots \sum_{q_n} \delta_{q_1+\cdots+q_n, q} P_X(q_1) \cdots P_X(q_n) s^q \quad (4.13)$$

$$= \sum_{n=0}^{\infty} P_N(n) \sum_{q=0}^{\infty} (\sum_{q_1} P_X(q_1) s^{q_1}) \cdots (\sum_{q_n} P_X(q_n) s^{q_n}) \delta_{q_1+\cdots+q_n, q} \quad (4.14)$$

$$= \sum_{n=0}^{\infty} (\sum_{q=0}^{\infty} P_X(q) s^q)^n \quad (4.15)$$

$$= \sum_{n=0}^{\infty} P_N(n) (g_X(s))^n \quad (4.16)$$

$$= g_N(g_X(s)) \quad (4.17)$$

which completes the proof.

We can now apply Eq. (4.9) to Eq. (4.3). If we by g_{Z_n} denote the generator function for the size of the population in the n -th generation, we notice that

$$g_{Z_1}(s) = \sum_{k=0}^{\infty} p_k s^k = g(s), \quad (4.18)$$

where $g(s)$ is the generator for the branching process itself with, we recall, p_k being the probability that a node has k out going branches. From the self-similarity of the branching process, it follows that the probability that $Z_n^{(k)} = q$, is equal to the probability that $Z_{n-1} = q$. We therefore have that $g_{Z_n^k}(s) = g_{Z_{n-1}}(s)$ and from Eq. (4.9) we derive the following iterative relationship:

$$g_{Z_n}(s) = g(g_{Z_{n-1}}(s)). \quad (4.19)$$

We can use this expression to analyse the probability for stopping or non-stopping. The probability x_n that the process has stopped at or before generation n is given by

$$x_n := \text{prob}\{Z_n = 0\} = g_{Z_n}(0). \quad (4.20)$$

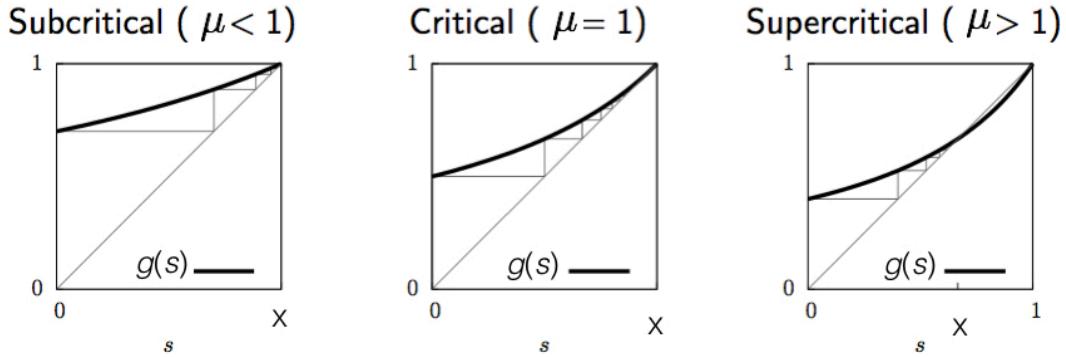


Figure 4.3: Roots of the Eq. $s = g(s)$ for different values of the average branching ration μ . (Fig. edited after fig. by Rosalba García Millán.)

Since

$$x_n = g_{Z_n}(0) = g(g_{Z_{n-1}}(0)) = g(x_{n-1}),$$

we have that $x_1 < x_2 < \dots < x_n < 1$ so the limit

$$x := \lim_{n \rightarrow \infty} x_n$$

will exist and it is straight forward by assuming the opposite and derive a contradiction that the limit will satisfy

$$x = g(x).$$

Therefore the probability that the process stops at some finite time is given by the roots of the equation $x = g(x)$. The value of the root x depends on how the generator function $g(s)$ behave at $s = 0$ and at $s = 1$. First we notice that $g(0) = p_0$, which we consider to be non-zero. Otherwise the process obviously never stops. Next we notice that a root $x < 1$ only exist if $g'(1) = \mu > 1$. See Fig. 4.3

We conclude that for $\mu \leq 1$ the process stops for some finite n with probability 1. Whereas, for $\mu > 1$ there is a probability x larger than 0 and less than 1 that the process stops at finite generation and hence a probability less than one that the process continues forever.

We are also now able to compute the average size generation n by use of the iterative relation in Eq. (4.19) and will show that

$$\langle Z_n \rangle = \mu^n. \quad (4.21)$$

I.e. exponential decay if $\mu < 1$ and exponential explosion if $\mu > 1$.

Proof

$$\begin{aligned} \langle Z_n \rangle &= g'_{Z_n}(1) = \frac{d}{ds} g(g_{Z_{n-1}}(s))|_{s=1} \\ &= [g'(g_{Z_{n-1}}(s)) \frac{d}{ds} g_{Z_{n-1}}(s)]|_{s=1} \\ &= [g'(g_{Z_{n-1}}(s)) g'(g_{Z_{n-2}}(s)) \frac{d}{ds} g_{Z_{n-2}}(s)]|_{s=1} \\ &= [g'(g_{Z_{n-1}}(s)) g'(g_{Z_{n-2}}(s)) \cdots g'(g_{Z_0}(s)) g'(s)]|_{s=1}. \end{aligned} \quad (4.22)$$

Now use that

$$g_{Z_n}(1) = g(g_{Z_{n-1}}(1)) = \cdots = g(g(\cdots(g(s))))|_{s=1} \quad (4.23)$$

and that $g(1) = 1$ and $g'(1) = \mu$ to conclude that

$$\langle Z_n \rangle = [g'(1)]^n = \mu^n. \quad (4.24)$$

4.1.1 Size of the progeny

We are also able to study the total progeny, or total size Y_n of a tree up to generation n ,

$$Y_n = 1 + Z_1 + Z_2 + \cdots + Z_n. \quad (4.25)$$

As we did above - see Fig. 4.2 we divide the process up according to the ancestor at generation 1 and make use of the self-similarity of the process. Let $Y_n^{(k)}$ denote the total number of nodes up to generation n , which originates from node k in the first generation. We have

$$Y_n = 1 + \sum_{k=1}^{Z_1} Y_n^{(k)}.$$

And therefore according to Eq. (4.9) we know that the generator functions for $\sum_{k=1}^{Z_1} Y_n^{(k)}$ is given by

$$g_{Y_n}(s) = g(g_{Y_{n-1}}(s)). \quad (4.26)$$

Note that a one line calculation shows that if $A = 1 + B$ where A and B are stochastic variables with $A, B \in \mathbb{N}$, then $g_A(s) = sg_B(s)$. We can therefore conclude that

$$g_{Y_n}(s) = sg(g_{Y_{n-1}}(s)). \quad (4.27)$$

From this relation it follows easily that

$$g_{Y_1} > g_{Y_2} > \cdots > g_{Y_n} \geq 0.$$

To see this, first observe that $sg(s) < s$ on $s \in [0, 1]$. Which can be seen from observing that with $f(s) = sg(s)$ we have $f'(0) = p_0 < 1$ and $f(1) = 1$. Next notice that $g_{Y_0}(s) = s$, since $Y_0 = 1$ with probability 1. Therefore

$$g_{Y_1}(s) = sg(s) < s = g_{Y_0}(s) \quad (4.28)$$

and then

$$g_{Y_2}(s) = sg(g_{Y_1}(s)) = sg(sg(s)) < sg(s) = g_{Y_1}(s). \quad (4.29)$$

an we can carry on like that.

The limit $g_{Y_\infty}(s) := \lim_{n \rightarrow \infty} g_{Y_n}(s)$ must exist and is given as a root in the equation

$$g_{Y_\infty}(s) = sg(g_{Y_\infty}(s)). \quad (4.30)$$

we can again graphically see, cf. Fig. 4.4, that there are two cases depending on the value of μ . Recall that the root $t_{root} = g_{Y_\infty}(s)$ and that

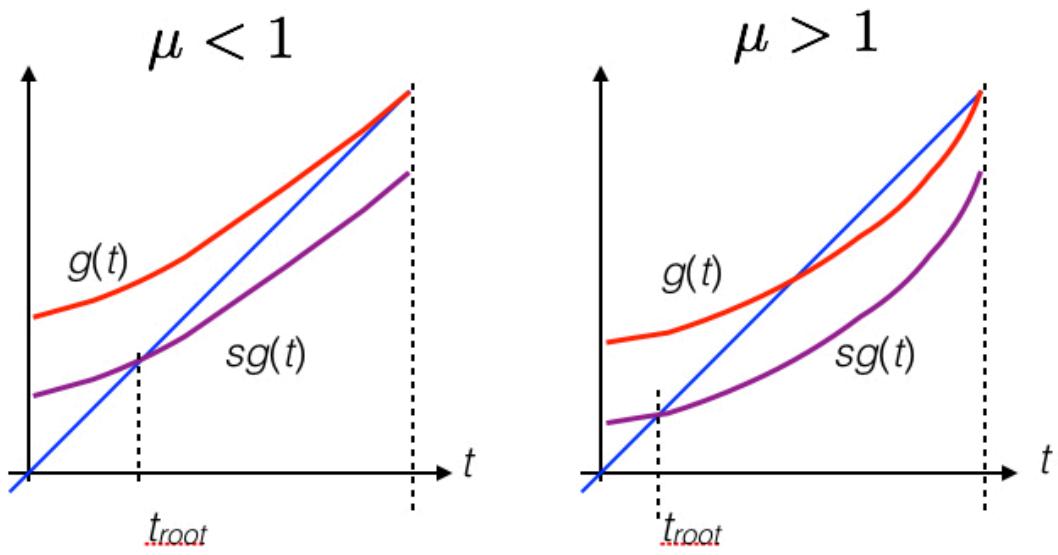


Figure 4.4: Roots of the Eq. (4.30) for different values of the average branching ratio μ .

$$g_{Y_\infty}(1) = \sum_{k=1}^{\infty} \text{Prob}\{Y_\infty = k\} \quad (4.31)$$

is the normalisation of the variable Y_∞ . For $\mu > 1$ we have $t_{root} < 1$ as $s \rightarrow 1$. Thus the stochastic variable Y_∞ is not normalised for $\mu > 1$. This corresponds to what we found above that for $\mu > 1$ the branching process can continue forever with a probability larger than zero, so a subset of the generated trees will have infinite size. This is in contrast to the situation for $\mu < 1$ where t_{root} approaches 1 as $s \rightarrow 1$. Corresponding to the fact that all trees stop for some finite n when $\mu < 1$. The two cases are separated by the critical value $\mu_c = 1$.

We saw above that for $\mu < \mu_c$, i.e. the subcritical case, the average generation size decays exponentially and for the supercritical case $\mu > \mu_c$ we have exponential growth of $\langle Z_n \rangle$ with n . This behaviour is also reflected in the average size of the progeny $\langle Y_\infty \rangle$, which is given by $g'_{Y_\infty}(1)$. We have

$$\begin{aligned} g_{Y_\infty}(s) &= sg(g_{Y_\infty}(s)) \Rightarrow \\ g'_{Y_\infty}(s) &= g(g_{Y_\infty}(s)) + sg'(g_{Y_\infty}(s))g'_{Y_\infty}(s) \Rightarrow \\ g'_{Y_\infty}(s) &= \frac{g(g_{Y_\infty}(s))}{1 - sg'(g_{Y_\infty}(s))}. \end{aligned} \quad (4.32)$$

Let $s = 1$ and for $\mu < 1$ make use of

$$g'(g_{Y_\infty}(1)) = g'(1) = \mu,$$

to conclude that

$$\langle Y_\infty \rangle = g'_{Y_\infty}(1) = \frac{g(1)}{1 - g'(1)} = \frac{1}{1 - \mu}. \quad (4.33)$$

So, as probably expected, $\langle Y_\infty \rangle \rightarrow \infty$ as $\mu \rightarrow 1$.

The probability that $Y_\infty = k$ can be found to behave as

$$\text{Prob}\{Y_\infty = k\} \sim k^{-3/2} e^{-k/k_0(\mu)}, \quad (4.34)$$

where $k_0(\mu) \sim (1 - \mu)^{-2}$. The standard references are

- T.E. Harris, *The Theory of Branching Processes*, Springer 1963.
- R. Otter, *The multiplicative process*, Annals of Math. Stat., **20**, 206-224, (1949).

4.1.2 Time to extinction

So far we have concentrated on the size of the trees Z_n , Y_n and Y_∞ . Next we turn to temporal aspects. Namely the time τ to extinction given by

$$\tau := \min\{n \geq 1 \mid Z_n = 0\}. \quad (4.35)$$

We see that τ is the generation number at which for the first time no offspring are present produced. In Fig. 4.1 $\tau = 6$.

We will sketch proofs of the following asymptotic behaviour for the accumulated distribution $P(\tau > n) := \text{prob}\{\tau > n\}$ for the branching processes with $\mu \leq 1$ and $\sigma < \infty$ ¹

$$\begin{aligned} \mu < 1 : P(\tau > n) &\sim \mu^n \text{ for large } n. \\ \mu = 1 : P(\tau > n) &\sim \frac{2}{\sigma^2 n} \text{ for large } n. \end{aligned} \quad (4.36)$$

Proof

Recall that the generator for the size of the n -th generation is $g_{Z_n}(s)$. And that the probability for stopping is

$$x_n = \text{prob}\{Z_n = 0\} = g_{Z_n}(0) = 1 - \text{prob}\{Z_n > 0\}$$

and therefore

$$P(\tau > n) = \text{Prob}\{Z_n > 0\} = 1 - g_{Z_n}(0).$$

For $\mu < 1$ we found that $x = \lim_{n \rightarrow \infty} x_n = 1$. So for large n we have $g_{Z_n}(0) \simeq 1$ and we can Taylor expand about $s = 1$. Below we do this while we also make use of Eq. (4.19) and the fact that $g(1) = 1$ and $g'(1) = \mu$.

$$\begin{aligned} P(\tau > n + 1) &= 1 - g_{Z_{n+1}}(0) = 1 - g(g_{Z_n}(0)) = 1 - g(1 - (1 - g_{Z_n}(0))) \\ &= 1 - [g(1) - (1 - g_{Z_n}(0))g'(1)] = 1 - [1 - (1 - g_{Z_n}(0))\mu] \\ &= \mu(1 - g_{Z_n}(0)) = \dots = \mu^n. \end{aligned} \quad (4.37)$$

For $\mu = 1$ we need to Taylor expand to 2nd order, while we make use of $g(1) = 1$ and

¹Here σ refers to the standard deviation of the offspring process described by the branching probabilities p_k .

$g'(1) = \mu = 1$. Hence,

$$\begin{aligned}
1 - g_{Z_{n+1}}(0) &= 1 - g(g_{Z_n}(0)) = 1 - g(1 - (1 - g_{Z_n}(0))) \\
&= 1 - [g(1) - g'(1)(1 - g_{Z_n}(0)) + \frac{1}{2}g''(1)(1 - g_{Z_n}(0))^2] \\
&\Downarrow \\
1 - g_{Z_n}(0) &= 1 - g_{Z_n}(0) - \frac{1}{2}g''(1)(1 - g_{Z_n}(0))^2.
\end{aligned} \tag{4.38}$$

We conclude that $x_n = 1 - g_{Z_n}(0)$ satisfies the iterative equation

$$x_{n+1} = x_n - bx_n^2, \tag{4.39}$$

with $b = g''(1)/2$. This is equation is solved asymptotically (i.e for $n \gg 1$) by $x_n = \alpha/n$. One see this for instance by substituting $x_n = \alpha/n$ into Eq. (4.39) and noticing that when $\alpha = 1/b$ the equation reduces to $n^2/b = n^2/b - 1/b$ so right and left hand is asymptotically equal. We conclude that $P(\tau > n) = 1 - g_{Z_n}(0) \sim 1/n$. In summary we have for probabilities

$$P(\tau = n) = \frac{d}{dn}(1 - P(\tau > n)) \sim \begin{cases} n\mu^{n-1} & \text{for } \mu \leq 1 \\ 1/n^2 & \text{for } \mu = 1. \end{cases} \tag{4.40}$$

We saw in Eq. (4.34) that the tree size distribution $\text{Prob}\{Y_\infty = k\}$ behave like a power law $k^{-3/2}$ multiplied by an exponential factor and that the latter goes to one as one when the average branching ratio $\mu \rightarrow 1$, to the critical value. The distribution for the first return time for a random walker is also described by the exponent $3/2$. Is this perhaps related?

Let us first recall what exactly is meant by the return time of a random walker. Think of a random walker with position $x(0)$ at time zero. See Fig. 4.5. Assume that $x(0) = 0$. Let, e.g, the walker increments Δ be distributed according to some distribution for which the first two moments exist and such that $\langle \Delta \rangle = 0$, so the walk is unbiased. We can consider the times $0 < T_1 < T_2 < \dots$ when the walker has returned to the origin for the first time T_1 , for the second time T_2 etc. , i.e. $x(T_1) = x(T_2) = \dots = 0$. The probability that $T_1 = T$ is (asymptotically) give by $T^{-3/2}$.

We can make a map between the branching process and the random walker in the following way. We think of walking on the tree generated by a branching process, for example the tree generated in Fig. 4.1. We enumerate the nodes by $i = 0, 1, 2, \dots$, see Fig. 4.6. Next we need to relate the increments of the walk to the off-spring

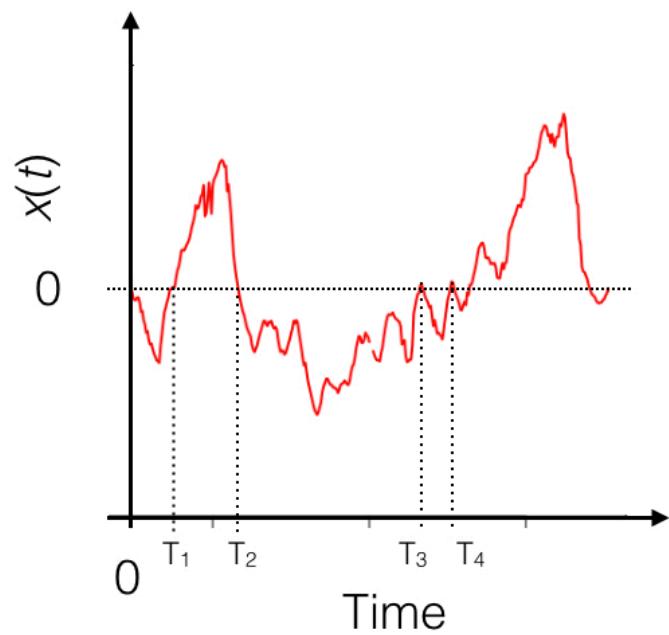


Figure 4.5: Walker moving up and down along the y-axis.

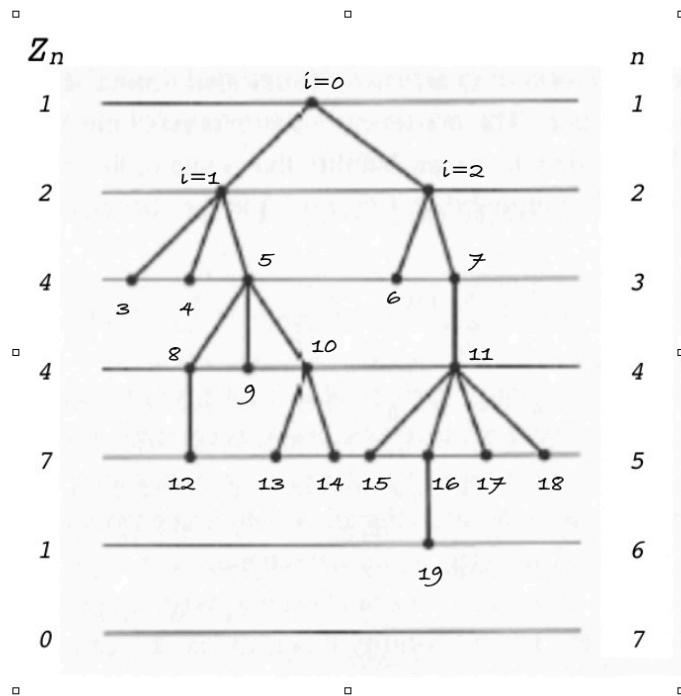


Figure 4.6: Walker visiting the nodes of a tree generated by a branching process.

production. We let the increment Δ_i be given by the number of off-spring emerging from a node, k_i minus one. And let the "position" x_i of our random walker be given by $x_{i+1} = x_i + \Delta_i$. I.e the random walk variable is updated according to

$$x_{i+1} = x_i + k_i - 1. \quad (4.41)$$

We can think of the walk in the following way. Whenever we arrive at a node the off-spring shooting off from that node is "revealed" to us. If we per definition assume $x_0 = 1$, after i steps, x_i is equal to the number of nodes revealed but not yet visited plus one. We of course increase the number of visited sites by one whenever we step on to new node. And if no off-spring sprouts from that node we lower x_i by one. But if $k > 0$ off-spring sprouts from the node that the number of non-visited sites increases by k and the net change to x_i is $k - 1$.

In Fig. 4.6 we have $x_0 = 1$, $x_1 = 1 + 2 - 1 = 2$, $x_2 = 2 + 3 - 1 = 4$, etc. Eventually we may be able to visit all the nodes, if new offspring are produced for a while at a rate lower than 1. When this happens the random walk stops and i is equal to $Y_\infty - 1$. For the random walk the value of i at which this happens corresponds to the first return of x_i to zero and its distribution is characterised by the exponent $3/2$, which then corresponds to distribution of Y_∞ being characterised by the same exponent.

Chapter 5

Networks: generator formalism

Networks play a central role in complexity science since it is often very natural to map a complex system to a network. Often the nodes will represent the components of the considered system, say individual human beings in a study of the sociology of some population, and the edges will correspond to relationships between the people involved. The edges could e.g. represent the flow of telephone calls. A directed edge from A to B would represent that A calls B but B never makes a call to A. Or one may use undirected edge if the it is of no interest or perhaps impossible to determine in which direction the influence points. One may use weighted edges, say the weight is proportional to the total minutes in a month A and B spend on the phone together or one may simply use a more binary approach and include an edge between A and B if they share calls and no edge if this never happens. There are many good courses and books on the theory of networks, e.g. the course in the Physics Department given by Tim Evans and Kim Christensen and the book *Networks. An Introduction* by M.E.J. Newman.

Here we will discuss how generator formalism can be useful when considering degree distributions and cluster size distributions. First some definitions. Consider an ensemble of networks for which the probability that a node selected with uniformly probability (i.e. same probability for all the nodes independent of whatever attributes they may have) has degree k^1 is given by p_k for $k = 0, 1, 2, \dots$. The generator for this degree distribution is given by

$$g_0(s) = \sum_{k=0}^{\infty} p_k s^k. \quad (5.1)$$

We next introduce the concept of *excess degree*. Select a node, call it α , of degree

¹Recall that by degree we mean number of edges attached to the node

$k \geq 1$. Follow one of the edges attached to node α in order to reach a neighbour node, call it β . Focus on the edges emerging from β different from the edge that took us from α to β . The number of such edges is called the excess degree of β . In Fig. 5.1 the degree of α is 5 and the excess degree of β is equal to three.

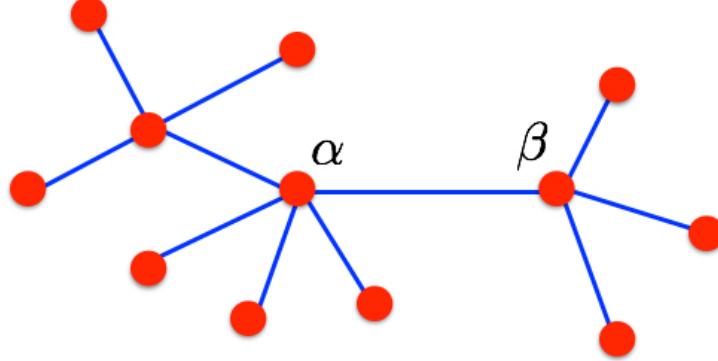


Figure 5.1: Walker visiting the nodes of a tree generated by a branching process.

Let q_k be the probability that the neighbour has excess degree k . The generator function for this distribution is

$$g_1(s) = \sum_{k=0}^{\infty} q_k s^k. \quad (5.2)$$

We will now derive a relation between $g_0(s)$ and $g_1(s)$. Recall that we consider random uncorrelated networks (called configuration models) with specific degree distributions. Assume that the total number of edges is E and the total number of nodes is N . The average degree is then given by $\langle k \rangle = E/N$. The expected number of nodes of degree k is Np_k , hence we have kNp_k edges sticking out from nodes of degree k . We call these for 1/2 edges. In total there is $2E$ half edges. Hence the probability that an edge we follow lands on a node of degree k is

$$\frac{kNp_k}{2E} = \frac{kp_k}{\langle k \rangle}. \quad (5.3)$$

The probability, q_k , that the node at the end of the edge has excess degree k - and therefore total degree $k + 1$ - is

$$q_k = \frac{(k+1)p_{k+1}}{\langle k \rangle},$$

and we can conclude that the generator for the excess degree distribution is given by

$$\begin{aligned}
g_1(s) &= \sum_{k=1}^{\infty} q_k s^k = \frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} (k+1) p_{k+1} s^k \\
&= \frac{1}{\langle k \rangle} \sum_{k=1}^{\infty} k p_k s^{k-1} = \frac{1}{\langle k \rangle} \frac{s}{ds} \sum_{k=0}^{\infty} p_k s^k \\
&= \frac{1}{\langle k \rangle} \frac{d}{ds} g_0(s),
\end{aligned} \tag{5.4}$$

and since $\langle k \rangle = g'_0(1)$ we conclude that

$$g_1(s) = \frac{g'_0(s)}{g'_0(1)}. \tag{5.5}$$

5.0.1 Cluster sizes

We now turn to a discussion of the cluster sizes in random networks. There are obvious similarity to the discussions in the section on percolation. We define

$$\pi_Q := \text{prob}\{\text{a randomly chosen node belongs to a (non - giant) cluster of size } Q\}. \tag{5.6}$$

The related generator function is

$$h_0(s) = \sum_{Q=1}^{\infty} \pi_Q s^Q. \tag{5.7}$$

Concerning normalisation we allow for giant cluster by assuming $\sum_Q \pi_Q = 1 - S$, where S is the strength of the giant cluster. Below we'll show that

$$h_0(1) = g_0(h_1(1)) = 1 - S,$$

with the generator function $h_1(s)$ to be defined in due course. First we demonstrate that small clusters are most likely of trees structure, i.e. unlikely to contain loops. To see this we compare the number of ways we can add a new edge to an existing cluster of size $Q \ll N$ nodes out of a total of N nodes. Assume the cluster is of tree structure. We can then compare the number of ways the new edge can be added such that it produces a loop. We attach one end of an edge to one of the Q nodes in the cluster. In order for this new edge to produce a loop the other end has to be connected to one of the other $Q - 1$ nodes in this cluster. This can be done in $Q - 1$ ways, which, for

large N is much smaller than the number of ways the new edge can be connected to a node that doesn't belong to the cluster, of which there are $N - Q$.

For this reason we will as a reasonable approximation neglect loops and concentrate on tree clusters. We analyse the distribution of cluster sizes by making use of their self-similar structure.

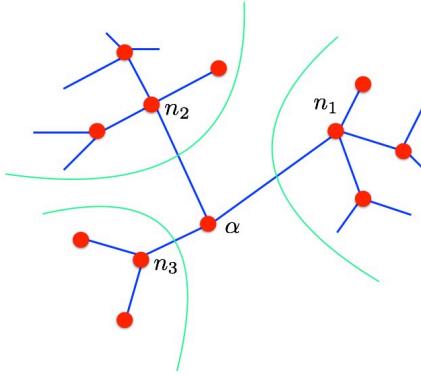


Figure 5.2: The size of the cluster containing node α is given by the sum of the size of the clusters originating from the neighbours of α .

We want to relate the size of the cluster containing a specific node α , see Fig. 5.2, to the size of the clusters originating from the neighbours of α . To do this we introduce the probability ρ_Q given by

$$\rho_Q = \text{prob}\{\text{node at end of edge belongs to a cluster of size } Q\} \quad (5.8)$$

and denote its generator by

$$h_1(s) = \sum_{Q=1}^{\infty} \rho_Q s^Q. \quad (5.9)$$

Let

$$P(S|k) = \text{prob}\{\text{the } k \text{ neighbours of node } \alpha \text{ belongs to a cluster with a total of } S \text{ nodes}\}, \quad (5.10)$$

we therefore have

$$P(S-1|k) = \text{prob}\{i \text{ belongs to a cluster of size } S\}. \quad (5.11)$$

This rather cumbersome set of definitions now allow us to observe that the probability that a arbitrarily chosen node α belongs to a cluster of size Q is given by

$$\pi_Q = \sum_{k=0}^{\infty} P(Q-1|k) p_k. \quad (5.12)$$

Where the factor p_k is the probability that the chosen node α happens to have degree k . We make use of the above to derive the following expression for the generator $h_0(s)$ for the cluster size distribution.

$$\begin{aligned}
h_0(s) &= \sum_{Q=1}^{\infty} \sum_{k=0}^{\infty} P(Q-1|k)p_k s^Q \\
&= s \sum_{Q=1}^{\infty} \sum_{k=0}^{\infty} P(Q-1|k)p_k s^{Q-1} \\
&= s \sum_{k=0}^{\infty} p_k \sum_{Q=0}^{\infty} P(Q|k)s^Q.
\end{aligned} \tag{5.13}$$

We observe that

$$\sum_{Q=0}^{\infty} P(S|k)s^Q$$

is the generator for the stochastic variable

$$Q = \sum_{n=1}^k Q_n.$$

which is the sum of the sizes of the sub-cluster the k neighbours belong to. Hence S is a variable of the form already considered in Eq. (4.7) and Eq. (4.9). In or case now the upper limit (the number of neighbours) is fixed at k , this corresponds to applying Eq. (4.9) to a case where the upper limit $N = N_0$ with probability 1. The generator for such a variable is $g_N(s) = s^{N_0}$. So for fixed upper limit Eq. (4.9) reduces to $g_{S_N}(s) = g_X(s)^{N_0}$. Now apply this to $Q = \sum_{n=1}^k Q_n$ keeping in mind that the generator for the sub-clustersizes Q_i of the neighbour clusters is $h_1(s)$ we obtain

$$\sum_{Q=0}^{\infty} P(Q|k)s^Q = [h_1]^k, \tag{5.14}$$

this result we substitute into Eq. (5.13) and get for the generator $h_0(s)$ of the probability that a randomly chosen node belongs to a cluster of a certain size

$$h_0(s) = s \sum_{Q=0}^{\infty} [h_1(s)]^Q = sg_0(h_1(s)), \tag{5.15}$$

where we made use of the fact that the generator for the degree distribution is denoted by $g_0(s)$.

Let us summarise. Eq. (5.15) relates the generator $h_0(s)$ for the probability that a randomly chosen node belongs to size Q cluster to the generator $g_0(s)$ for the degree distribution and the generator $h_1(s)$ for the probability that the size of the cluster at the end of an edge has a size Q_i . We are going to use this relation to discuss the existence of a giant cluster.

5.0.2 The giant cluster

Similar to our discussion in percolation, we can investigate the conditions for a *giant cluster* forms. By that we mean a cluster of a size that grows, or scales as it is typically called, with the total number of nodes N . I.e. this is cluster that keeps growing as we increase N , similar to the system spanning cluster we discussed in relation to percolation.

Since $h_0(s)$ is the generator for the probability that a randomly chosen nodes belongs to a cluster of some specific finite size $Q = 1, 2, 3, \dots < \infty$, we have that $h_0(1)$ is the total probability that a random node belongs to some finite cluster. The complementary to belonging to some finite cluster is to belong to the giant cluster. If we let the strength of the giant cluster be denoted by S , we have according to Eq. (5.15) that

$$1 - S = h_0(1) = g_0(h_1(1)) \quad (5.16)$$

i.e.

$$S = 1 - g_0(h_1(1)) \quad (5.17)$$

Hence we have to determine $h_1(s)$. We recall that $h_1(s)$ is the generator for the probability that the node at the end of the edge is of a specific finite size Q_i . This means that

$$\begin{aligned} h_1(1) &= \text{prob}\{\text{node not connected to giant through any neighbour}\} \\ &= \text{prob}\{\text{each neighbour is not connected to giant}\}. \end{aligned} \quad (5.18)$$

Assume that the specific neighbour has excess degree k , which happens with probability q_k . To obtain the probability the probability that each of the neighbours are not connected to the giant we have to sum over all possible excess degrees, i.e.

$$h_1(1) = \sum_{k=0}^{\infty} q_k [h_1(1)]^k = g_1(h_1(1)). \quad (5.19)$$

Where we used that $g_1(s)$ is the generator function for the excess degree. Eq. (5.19) determines $h_1(1)$ as the smallest root² in the equation

$$x = g_1(x). \quad (5.20)$$

Example

Consider an exponential degree distribution

$$p_k = (1 - e^{-\lambda})e^{-\lambda k}, \text{ with } \lambda > 0. \quad (5.21)$$

The generator for this degree distribution is

$$\begin{aligned} g_0(s) &= \sum_{k=0}^{\infty} p_k s^k = (1 - e^{-\lambda}) \sum_{k=0}^{\infty} e^{\lambda k} s^k \\ &= (1 - e^{-\lambda}) \sum_{k=0}^{\infty} \left[\frac{s}{e^{\lambda}} \right]^k = \frac{1 - e^{-\lambda}}{1 - se^{-\lambda}} = \frac{e^{\lambda} - 1}{e^{\lambda} - s}. \end{aligned} \quad (5.22)$$

The generator for the excess degree is according to Eq. (5.5)

$$g_1(s) = \frac{g'_0(s)}{g'_0(1)} = \left(\frac{e^{\lambda} - 1}{e^{\lambda} - s} \right)^2. \quad (5.23)$$

According to Eq. (5.20) we need to find the root of the following Eq.

$$x = \left(\frac{e^{\lambda} - 1}{e^{\lambda} - x} \right)^2. \quad (5.24)$$

Note $x = 1$ is a trivial solution, and factorise the equation to obtain

$$x^2 - (2e^{\lambda} - 1)x + (e^{\lambda} - 1)^2 = 0. \quad (5.25)$$

The root gives us $h_1(1)$ and from Eq. (5.17) we obtain

$$S = 1 - g_0(h_1(1)) = 3/2 - \sqrt{e^{\lambda} - 3/4}.$$

Fig. 5.3 shows how S depends on λ . There is no giant cluster for $\lambda > \ln 3$: the degree distribution decays too fast. For $0 < \lambda < \ln 3$ a giant cluster exists and for $\lambda = 0$ all nodes belong to the giant.

²Note $x = 1$ is always a root

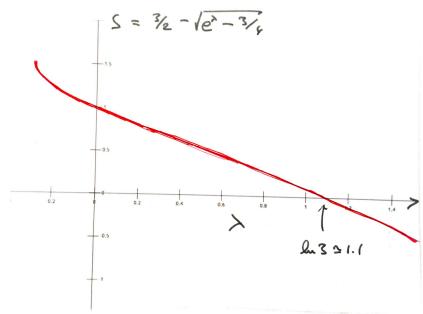


Figure 5.3: The strength of the giant cluster as function of λ .

Chapter 6

Intermittency: Extreme Value Statistics, Record Dynamics and Tangent Map

As witnessed by the famous *Guinness Book of Records*, record events of all kinds have always attracted a great deal of interest. Aside from their popular appeal, extremes and their properties are important topics of mathematical statistics [48, 43]. Furthermore, the behaviour of complex systems is often determined by the unusual rather than the average or typical events and in some contexts, record sized events have important implications, e.g. in population dynamics, mutants surpassing the current standard for highest reproductive success contribute disproportionately to the genetic pool of the next generation. Record events might therefore be of significance to biological evolution, as in particular suggested by the catchphrase ‘survival of the fittest’¹.

Darwinism’s impact on biology—and human culture at large—can hardly be exaggerated. In recent years computer optimization methods have appeared, e.g. Genetic Algorithms [25] and Extremal Optimization[9], which are inspired by evolution dynamics and which, implicitly or explicitly, utilize records. In a computer program, or in a social context, the function, entity or organization in charge of memorizing the existing record and detecting its eventual obliteration is clearly distinct from the process being monitored. By contrast, record keeping in biological evolution is fully integrated in the dynamics. In cases where record fluctuations act as seeds for an amplifying dynamical process, the magnitude of the seed bears no direct relation to the

¹The phrase is attributed to Herbert Spencer in Principles of Biology, 1864. The underlying concept can be very problematic. The notion of *the fit* or *the fittest* has, as is well known, been misused in the extreme during the 20th century.

magnitude of the final outcome of the amplification, e.g. a snow avalanche can be triggered by sound vibrations above a certain threshold. However, its impact is more related to the amount of snow on the ground and to the slope and shape of the valley than to the intensity of the sound.

How record sized events in a stochastic time series may be related to, and indeed be a determining factor for, the dynamics of complex systems is discussed in some detail in Sibani and Jensen, *Stochastic Dynamics of Complex Systems*. In the following we focus on the temporal statistics of records: How records within a stationary and δ -correlated time series are statistically distributed in time. Secondly, we analyse records in a time series which is both non-stationary and correlated: The position of a random walker moving along a line.

We first we consider the distinct issue of extremal value statistics, i.e. the distribution of the largest of n observations.

6.1 Extremal Value Statistics

Record and extreme value statistics both deal with large events, but the former is concerned with the *temporal* distribution of records in a time ordered series, while the latter is concerned with the *size* distribution of the largest element in a set of n random variables. Record statistics implies an ordering since a record refers to what has happened earlier: The world record for long jump was broken today because Mr. Tall managed to jump 4 mm longer than the previous record set by Mr. Short two years ago. Thousands of long jumps have been undertaken between those by Mr. Short and Mr. Tall. But none of these released a new record because they were all shorter than Mr Short's. The focus of record statistics, at least in the context of our book and as discussed in the preceding sections, is to understand the statistics of the time instances at which the considered quantity achieves a value higher than any previous recorded value. One is less interested in the actual value of the record, here the length of the jump.

Addressing the long jump from the viewpoint of extreme value statistics, the question could e.g. be the following. One hundred equally competent long jumpers gather to practice their passion for long jumps. We have been hired to construct the jump pit. Now the question arises: How long do we need to make the pit in order to make sure that the longest, i.e. the extreme, of the 100 jumps doesn't reach beyond the soft sand of the pit. Here we don't care about the sequential ordering of the 100 jumps, but only about the size of the extreme jump. Since if we ensure that the length of the pit is longer by a safety margin than the statistically expected extreme, all jumpers are safe. To do the statistics we obviously need information about the distribution of the

length of the individual jumps.

Let us now develop the formalism used to treat such issues mathematically. Consider a sequence $x(t_i)$ of identically distributed and independent real valued stochastic variables for $t_i = 1, 2, 3, \dots$. A record happens at time t_r if and only if

$$x(t_r) > x(t) \text{ for all } t < t_r. \quad (6.1)$$

As discussed in Section 6.2, the sequence of records falling between times t_1 and t_2 is to a good approximation a Poisson process whose average is proportional to $\ln(t_2/t_1)$. Remarkably, this result is largely independent of the distribution of $x(t)$. In contrast extreme value statistics is concerned with the functional form of the distribution of the greatest (or smallest) values of the x 's. As one would expect, this distribution turns out to depend on the distribution of the individual variables.

When studying extreme value statistics we consider the set $S_n = \{x_1, x_2, x_3, \dots, x_n\}$, where $x_i \stackrel{\text{def}}{=} x(t_i)$, and discuss how the largest value²

$$\chi_n = \max[S_n] \quad (6.2)$$

is distributed in the limit of large n . Our heuristic discussion aims at explaining why there are three different classes of asymptotic extreme value distributions and only one asymptote for record statistics. More in depth treatment can be found in Refs. [28, 42, 24, 50].

Let $P(x_i)$ denote the common probability density function (PDF) of the x_i 's. We furthermore introduce the cumulative distributions

$$F_<(x) = \int_{-\infty}^x dx' P(x') \quad (6.3)$$

and $F_>(x) = 1 - F_<(x)$. Since the variables x_1, x_2, \dots are assumed to be independent,

$$\begin{aligned} H_n(x) &\stackrel{\text{def}}{=} \text{Prob}\{\chi_n < x\} \\ &= \text{Prob}\{x_1 < x, x_2 < x, \dots, x_n < x\} \\ &= \text{Prob}\{x_1 < x\} \text{Prob}\{x_2 < x\} \cdots \text{Prob}\{x_n < x\} \\ &= [F_<(x)]^n. \end{aligned} \quad (6.4)$$

Since the cumulative distribution satisfies $F_<(x) \leq 1$, Eq. 6.4 implies that for fixed x and in the limit $n \rightarrow \infty$ the distribution $\lim_{n \rightarrow \infty} H_n(x)$ approaches either zero or

²The smallest would also do, since a change of sign $x_k \mapsto -x_k$ maps the smallest into the largest value.

one depending on whether $F_<(x) < 1$ or $F_<(x) = 1$. This indicates that the limit $n \rightarrow \infty$ must be treated with care. Mathematical details can be found in the literature e.g. [28, 42, 24, 50]. As the number of elements in the sampling set S_n increases, so will typically their maximum χ_n . This suggests introducing a linear transformation $\chi_n \mapsto (\chi_n - b_n)/a_n$ in a manner similar to what is done to a sum of random variables $S_N = \sum_{i=1}^N x_i$, each term having the mean μ and variance σ^2 . In that case, $(S_N - \mu N)/(\sqrt{N}\sigma)$ converges to a normal distribution of zero average and unit variance.

Does a limiting distribution $H(x)$ similarly exist for $(\chi_n - b_n)/a_n$ and can we obtain some insight into its properties? The answer is yes, based on the following heuristic argument which goes back to the 1920's [23, 22]. The idea is based on deriving an equation for the distribution $H_n(a_n x + b_n)$ of the appropriately scaled variable in the limit of large n . The main property used is that the largest value of a union of sets of numbers equals the largest of the maxima of each of the sets.

Consider then N sets each containing n realizations of the stochastic variable x , namely:

$$\{x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}\}, \{x_1^{(2)}, x_2^{(2)}, \dots, x_n^{(2)}\}, \dots, \{x_1^{(N)}, x_2^{(N)}, \dots, x_n^{(N)}\}. \quad (6.5)$$

Recalling that $H_n(x)$ is the (cumulative) distribution of the greatest value χ_n , $H_N(a_N x + b_N)$ must then be the distribution function of the greatest value of the combined set

$$S_{Nn} = \{x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}, x_1^{(2)}, x_2^{(2)}, \dots, x_n^{(2)}, \dots, x_1^{(N)}, x_2^{(N)}, \dots, x_n^{(N)}\}. \quad (6.6)$$

with appropriate choice of a_N and b_N . But since

$$\begin{aligned} \chi_{Nn} &= \max[S_{Nn}] \\ &= \max[\{\max[\{x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}\}], \dots, \max[\{x_1^{(N)}, x_2^{(N)}, \dots, x_n^{(N)}\}]\}] \end{aligned}$$

we arrive at the so-called stability postulate [23, 22], which reads

$$H_N(a_N x + b_N) = [H_n(x)]^N. \quad (6.7)$$

If we boldly assume that for sufficiently large N and n the functions H_N and H_n can be replaced in Eq. 6.7 by their limit distribution $H = \lim_{m \rightarrow \infty} H_m$, we conclude that

$$H(a_N x + b_N) = [H(x)]^N. \quad (6.8)$$

As it turns out, one obtains three different types of extreme value asymptotes depending on the choices of the transformation parameters a_N and b_N . By this we mean that

$$\text{Prob}\{X_N = (a_N \chi_N + b_N) < x\} \rightarrow H(x) \text{ for } N \rightarrow \infty. \quad (6.9)$$

for appropriate choice of a_N and b_N . Note the similarity to the situation for the sum of iid numbers and the usual central limit theorem.

The choice $a_N = 1$ gives the Gumbel type, see Eq. 6.12 below. The two other types correspond to $a_N \neq 1$. These are called Fréchet, see Eq. 6.16 and Weibull, see Eq. 6.18, respectively [28, 42, 24, 50]. Before listing all three types, let us briefly discuss how Eq. 6.7 allows one to study the functional form of the limit. We concentrate on the first case where $a_N = 1$ and where the stability condition, Eq. 6.7, reads $[H(x)]^N = H(x + b_N)$. Next we apply Eq. 6.7 to $H(x + b_N)$ and get

$$[H(x)]^{Nn} = [H(x + b_n)]^N = H(x + b_N + b_n). \quad (6.10)$$

However, considering that $[H(x)]^{Nn} = H(x + b_{Nm})$ also holds,

$$b_N + b_n = b_{Nm}. \quad (6.11)$$

This implies that $b_N = \alpha \log(N)$, where α is a constant. If we choose $\alpha = -1$, Eq. 6.7 is satisfied by $H(x) = \exp(-\exp(-x))$ since then

$$H^N(x) = \exp(-N \exp(-x))$$

and

$$H(x - \log(N)) = \exp(-\exp(-x + \log(N))) = \exp(-N \exp(-x)) = H^N(x).$$

The three different types of asymptotic distributions are given by:

1. Gumbel type

$$H(x) = \exp[-\exp(-x)] \quad (6.12)$$

where $-\infty < x < \infty$. The Gumbel limit applies when the probability density function $P(x)$ for the stochastic variables x_i vanishes sufficiently fast for large x , (see [24]). More precisely, when for some finite value a , the integral

$$\int_a^{\omega(F)} (1 - F(x))dx \quad (6.13)$$

is finite. The upper limit of the integral is given by $\omega(F) = \sup\{x : F(x) < 1\}$. Let $\alpha(F) \stackrel{\text{def}}{=} \inf\{x : F(x) > 0\}$ and for $\alpha(F) < t < \omega(F)$ let

$$R(t) = \frac{\int_t^{\omega(F)} (1 - F(x))dx}{1 - F(t)}. \quad (6.14)$$

The form in Eq. 6.12 applies when for all $x \in \mathbb{R}$

$$\lim_{t \rightarrow \omega(F)} \frac{1 - F(t + xR(t))}{1 - F(t)} = e^{-x}. \quad (6.15)$$

2. Fréchet type

$$H(x) = \begin{cases} \exp(-x^{-\gamma}) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases} \quad (6.16)$$

The exponent is determined from the limit

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\gamma}. \quad (6.17)$$

3. Weibull type

$$H(x) = \begin{cases} 1 & \text{if } x > 0 \\ \exp[-(-x)^\gamma] & \text{if } x \leq 0. \end{cases} \quad (6.18)$$

The exponent is in this case determined by the limit

$$\lim_{t \rightarrow \infty} \frac{1 - F(\omega(F) - \frac{1}{tx})}{1 - F(\omega(F) - 1/t)} = x^{-\gamma}. \quad (6.19)$$

Here ‘the end point’ of the distribution $F(x)$, given by $\omega(F) = \sup\{x : F(x) < 1\}$, is assumed to be finite.

The normalization coefficients a_n and b_n are not unique and possible choices are discussed in [24] and [28]. Furthermore, the range of attraction of the three different types of asymptotes is still a matter of research. Rather than entering into these matters, we will just consider some typical examples.

Exponential probability density Assume the probability density function $f(x)$ for the individual stochastic variables x_i to be given by

$$f(x) = \begin{cases} e^{-x} & \text{for } x \leq 0 \\ 0 & \text{for } x < 0. \end{cases} \quad (6.20)$$

The corresponding cumulative distribution function is $F(x) = \int_0^x e^{-x'} dx' = 1 - e^{-x}$. The condition described in Eq. 6.13 is fulfilled since $\omega(F) = \infty$ and

$$\int_0^\infty [1 - F(x)] dx = \int_0^\infty e^{-x} dx = 1 < \infty.$$

The quantity $R(t)$ in Eq. 6.14 becomes

$$R(t) = \frac{\int_t^\infty e^{-x} dx}{e^{-t}} = 1.$$

Hence the exponential PDF falls within the Gumbel class.

Power-law probability density Assume that the probability density function, or PDF, is given by

$$f(x) = \begin{cases} (1-a)x^{-a} & \text{for } x \geq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (6.21)$$

We assume $a > 1$ to ensure normalizability, i.e. the convergence of $\int_1^\infty f(x)dx$. The cumulative distribution

$$F(x) = \begin{cases} 0 & \text{for } x < 1 \\ \int_1^x (1-a)t^{-a} dt = 1 - x^{1-a} & \text{for } x \geq 1, \end{cases} \quad (6.22)$$

is characterized by $\omega(F) = \infty$ and fulfills

$$\frac{1 - F(tx)}{1 - F(t)} = \frac{1 - (1 - (tx)^{1-a})}{1 - (1 - t^{1-a})} = x^{-(a-1)}, \quad (6.23)$$

i.e. the assumed power-law distribution corresponds to the Fréchet case with $\gamma = a - 1$.

Bounded probability density Next, as an example of a PDF on a bounded interval consider

$$f(x) = \begin{cases} 1 & \text{for } x \in [0, 1] \\ 0 & \text{otherwise.} \end{cases} \quad (6.24)$$

The cumulative distribution is given by

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \in [0, 1] \\ 1 & \text{for } x \geq 1 \end{cases} \quad (6.25)$$

We have $\omega(F) = 1$, and to apply Eq. 6.19 we need

$$F^*(x) = F(1 - 1/x) = \begin{cases} 1 & \text{for } x \leq 0 \\ 0 & \text{for } 0 < x < 1 \\ 1 - 1/x & \text{for } x \geq 1. \end{cases} \quad (6.26)$$

and hence for large t

$$\frac{1 - F^*(tx)}{1 - F^*(t)} = \frac{1 - (1 - \frac{1}{tx})}{1 - (1 - \frac{1}{t})} = \frac{1}{x}. \quad (6.27)$$

From which we conclude that we are dealing with a Weibull or type 3 case with $\gamma = 1$ in Eq. 6.19.

6.2 Records in White Noise

Record sized entries within a stationary series of statistically independent and identically distributed random numbers, so-called *white noise*, form a subseries whose interesting statistical properties follow from simple arguments [48, 68, 66, 50], e.g. since it gets increasingly harder for an entry to qualify as the next record, records appear at a decreasing rate and the subseries of record sized entries is clearly not stationary. When the stationary time series from which the records are extracted is *white noise* the number of records in the time interval $(1, t)$ is to a good approximation Poisson distributed with average $\ln t$. Equivalently, if the k 'th record appears at times t_k , the ratios $\ln(t_k/t_{k-1})$ are independent random variables with an exponential distribution.

Consider first an ordered sequence of $l + 1$ random numbers drawn independently from a common distribution, and indexed from 0 to l . We exclude in the following distributions supported on a finite set, as these eventually lead to a record which can never be beaten. Apart from this modest constraint, the form of the distribution is immaterial. By definition, the first entry of the sequence ($k = 0$) is a record. Other entries are records if and only if they are larger than all their predecessors.

Next, denote by $P_n(l, [0, m_1 \dots m_{n-1}])$ the probability that n records are located at entries $0, m_1, \dots, m_{n-1}$, and by $p_n(l)$ the probability that n out of the $l + 1$ entries are records, regardless of their location. Clearly, in the case $n = 1$, the first entry must be the largest. As the largest entry can be anywhere with equal probability,

$$P_1(l, [0]) = p_1(l) = \frac{1}{l+1}, \quad l = 0, 1 \dots \infty. \quad (6.28)$$

For two records, one occurring at 0 and the other at $m_1 \leq l$, we find

$$P_2(l, [0, m_1]) = \frac{1}{m_1} \frac{1}{l+1}, \quad 1 \leq m_1 \leq l. \quad (6.29)$$

The first factor on the r.h.s. of the equation is the probability that the zero'th entry is the largest of the first m_1 entries, and the second is the probability that the largest

entry of the whole series be located at m_1 . Clearly,

$$p_2(l) = \frac{1}{l+1} \sum_{m_1=1}^l \frac{1}{m_1} = \frac{H_l}{l+1} \quad (6.30)$$

where the *harmonic number* $H_l = \sum_{k=1}^l 1/k$ satisfies

$$H_l = \ln(l+1) + \gamma + \mathcal{O}(l^{-1}) \quad (6.31)$$

with $\gamma = 0.57721\dots$ being the Euler–Mascheroni constant. Turning now to a generic n , we find, by the same arguments,

$$P_n(l, [0, m_1, m_2, \dots, m_{n-1}]) = \frac{1}{l+1} \prod_{k=1}^{n-1} \frac{1}{m_k}, \quad (6.32)$$

with; $1 \leq m_1; m_{k-1} < m_k; m_{n-1} \leq l$.

To obtain $p_n(l)$, the distribution $P_n(l, [0, m_1, m_2, \dots, m_{n-1}])$ must be summed over all possible values of $m_1, m_2 \dots m_{n-1}$. Unfortunately, this rather cumbersome operation does not lead to a closed form expression. A closed form expression can however be reached in the continuum limit, where all sums turn into integrals. To this end, let δt be a small time interval, and assume that a time series is produced by sampling a stationary noise signal at regular intervals of duration δt . Furthermore, let $t_k \stackrel{\text{def}}{=} m_k \delta t$ and $t \stackrel{\text{def}}{=} l \delta t$ be the time of occurrence of the k 'th record in the series, and the total observation time, respectively. Unless the noise is δ -correlated, contiguous values of the time series will of course become correlated in the limit $\delta t \rightarrow 0$. To avoid correlations, we now assume that the noise is white. Noting that $m_k = t_k / \delta t$, and taking the limit $\delta t \rightarrow 0$, we find

$$p_n(t) = \frac{1}{t} \int_{t_{n-2}}^t \frac{dt_{n-1}}{t_{n-1}} \prod_{k=2}^{n-2} \int_{t_{k-1}}^{t_{k+1}} \frac{dt_k}{t_k} \int_1^{t_2} \frac{dt_1}{t_1}. \quad (6.33)$$

As the integrand in the above expression is a symmetric function of the arguments t_1, \dots, t_{n-1} , any permutation of the arguments does not change the integral. Summing over all permutations of the order of the $n - 1$ variables and dividing by $(n - 1)!$ leaves the expression unchanged and yields

$$p_n(t) = \frac{1}{t} \frac{1}{(n-1)!} \left(\int_1^t \frac{dz}{z} \right)^{n-1} = \frac{1}{t} \frac{\log(t)^{n-1}}{(n-1)!}, n = 1, 2, \dots \infty \quad (6.34)$$

The expression can be recognized as a Poisson distribution with expectation value

$$E(n) = \log(t). \quad (6.35)$$

Clearly, the number of records falling between times t_w and $t > t_w$ is the difference of two Poisson processes, and hence itself a Poisson process with expectation value $E(n) = \ln(t) - \ln(t_w) = \ln(t/t_w)$. The average number of records per unit of time decays as

$$\frac{dE(n)}{dt} = \frac{1}{t}. \quad (6.36)$$

A third property of Poisson distributions, namely the exponential distribution of the waiting time between successive events, leads in connection with Eq. 6.34 to the following: Denoting by $\tau_k = \ln t_k$ the logarithm of the time at which the k 'th record occurs, the stochastic variables $\Delta_k = \tau_{k+1} - \tau_k = \ln(t_{k+1}/t_k)$, $k \geq 1$ and $\Delta_0 = \tau_1$ are statistically independent and exponentially distributed with unit average:

$$P_{\Delta_k}(x) = \text{Prob}(\Delta_k < x) = 1 - \exp(-x). \quad (6.37)$$

Since

$$\tau_k = \sum_{i=0}^{k-1} \Delta_i \quad (6.38)$$

is a sum of k independent exponential variables with unit average, it follows that τ_k is Gamma distributed, with density given by

$$P_{\tau_k}(t) = \frac{t^{k-1}}{(k-1)!} e^{-t}. \quad (6.39)$$

Both the average and the variance of the distribution are equal to k . By Jensen's inequality [63] we find

$$\ln(E(t_k)) \geq E(\ln(t_k)) = k. \quad (6.40)$$

For large k , the Gamma distribution can be approximated by a Gaussian distribution, and the waiting time t_k is therefore approximately log-normal.

A simple generalization of the results just obtained replaces the stationary time series by α mutually independent and stationary time series. For each of these, records are defined as discussed. The total number of records observed in a given time interval is then the sum of α independent Poisson variables, and hence itself a Poisson process. The expected number of records falling between times $t_1 < t$ and t is therefore

$$E(n) = \alpha \ln(t/t_1), \quad (6.41)$$

and the full distribution is

$$p_n([t_1, t]) = \left(\frac{t}{t_1} \right)^{-\alpha} \frac{(\alpha \ln(t/t_1))^{n-1}}{(n-1)!}, \quad n = 1, 2, \dots \infty; \quad \alpha = 1, 2, \dots \infty. \quad (6.42)$$

To summarize, the number of records occurring between times t_1 and $t > t_1$ in a time series of stationary white noise is a Poisson process. The expectation value of the process is proportional to $\ln(t/t_1)$ *irrespective* of how the noise values are distributed. This result changes when the assumptions on the underlying time series are relaxed, see e.g. Ref. [43] for a discussion of non-stationary series. Below we consider a different case, where the series is correlated and non-stationary.

6.3 Records and First Passage in Brownian Motion

The positions $x(t_0), x(t_1), \dots$ of a random walker moving in an infinite one-dimensional lattice $\{\dots -n, -n+1, \dots 0, \dots n, n+1, \dots\}$ constitute a time series which is neither δ -correlated nor stationary. As we shall see, the statistical properties of the records in this type of time series are rather different from the case treated in the previous section.

Since the random walk problem is translationally invariant, we can assume, without loss of generality, that the walker is initially located at the origin, i.e. $x(t=0) = 0$. As usual, the $t=0$ entry counts as the first record, while subsequent records occur at times larger than their predecessors. The time T_1 elapsed between the first and the second record coincides with the time of *first passage* through position $x=1$. In general, the time T_k elapsed from the k 'th to the $k+1$ 'th record coincides with the time of first passage through position $x=k$ for a walker starting at $x=k-1$. The T_k 's are mutually independent and identically distributed stochastic variables. We can hence use the symbol T for the time of first passage to a site which is the right neighbour of the initial position. The walker remains at its current position for a time T_r which is random and has an exponential distribution of unit average. Jumps are to the right or to the left with probabilities r_r and $1 - r_r = r_l$, respectively.

Let $f_{k,l}(t)$ be the probability density function (PDF) that the first passage through position k with start point at l occurs at time t . Since, as mentioned, this PDF only depends on $k-l$, we use $f_T(t)$ for any $f_{k,k-1}(t)$. Secondly, let $P_n(t)$ be the probability that precisely n records occur in the interval $[0, t]$. Clearly, $n > 1$ records in time t require $n-1$ records at an earlier time $t' < t$. A single record at time t implies that the first passage through 1 is never achieved. In summary,

$$P_1(t) = 1 - \int_0^t f_T(t') dt' \quad (6.43)$$

$$P_n(t) = \int_0^t P_{n-1}(t') f_T(t-t') dt', \quad n = 2, 3, \dots \quad (6.44)$$

Similar arguments lead to an integral equation for f_T : The walker can jump to the right and hit site k at his first move. If the first jump is to the left, the walker has to return to his original position (zero) at time $t' < t$ before a renewed attempt to hit site 1 can be made. The corresponding PDF is denoted by $f_{k,k-2}$, and

$$f_T(t) = r_r e^{-t} + r_l \int_0^t e^{-t'} f_{k,k-2}(t - t') dt', \quad (6.45)$$

where r_r and r_l are the probabilities of jumping to the right and left, respectively. Furthermore, we note that

$$f_{k,k-2}(t) = \int_0^t f_T(t') f_T(t - t') dt' \quad (6.46)$$

is itself a convolution. All the equations above involve convolutions and can be solved using Laplace transforms which map convolutions into products. Using a tilde for Laplace transformed functions we obtain

$$\tilde{P}_1(s) = \frac{1 - \tilde{f}_T(s)}{s} \quad (6.47)$$

$$\tilde{P}_n(s) = \tilde{P}_{n-1}(s) \tilde{f}_T(s) \quad n = 2, 3, \dots, \quad (6.48)$$

and

$$\tilde{f}_T(s) = \frac{r_r}{(1+s)} + \frac{r_l}{(1+s)} \tilde{f}_T^2(s). \quad (6.49)$$

Eqs.6.47 and 6.48 are solved by

$$\tilde{P}_n(s) = \frac{\tilde{f}_T^{n-1}(s) - \tilde{f}_T^n(s)}{s} \quad n = 1, 2, \dots \quad (6.50)$$

Turning now to f_T , we note that the physically relevant solution of Eq. 6.49 must fulfil $0 < f_T(s) \leq 1$ for all s . We note that $\tilde{f}_T(s=0) = \int_0^\infty f_T(t') dt'$ is the probability that the position $x = 1$ will be visited at all. As it turns out, for $r_l > r_r$ one has $\tilde{f}_T(s=0) < 1$. The walker has in this case a non-zero probability of escaping to $-\infty$, without ever visiting $x = 1$. Correspondingly, there is a non-zero probability that the initial record will never be beaten. Intuitively, this is due to a net drift to the left. In the extreme case $r_l = 1$, the walker moves at constant negative velocity, and $x = 1$ can never be visited. For $r_r > r_l$ the solution of Eq.6.49 is:

$$\tilde{f}_T(s) = \frac{1+s}{2r_l} \left(1 - \left(1 - \frac{4r_r r_l}{(1+s)^2} \right)^{1/2} \right), \quad r_r > r_l. \quad (6.51)$$

Let us first consider the interesting case $r_l = r_r = 1/2$. There is no net average velocity, and, as we shall see, the time elapsed between successive records has infinite expectation value. A calculation gives

$$\tilde{f}_T(s) = \frac{2(1+s) - 2(s(2+s))^{1/2}}{2} = \frac{s^{1/2} - (2+s)^{1/2}}{s^{1/2} + (2+s)^{1/2}}. \quad (6.52)$$

Inverting the Laplace transform yields

$$f_T(t) = \frac{e^{-t}}{t} I_1(t), \quad (6.53)$$

where I_1 is the modified Bessel function of order one. Using a different route, the same result is derived by Feller, see Eq. (17.13) in Ref. [21]. We note that

$$f_T(t) \rightarrow \frac{t^{-3/2}}{(2\pi)^{1/2}} \quad \text{for } t \rightarrow \infty. \quad (6.54)$$

Hence, f_T has no finite average, as anticipated. The Laplace transform of the expected number of record-breaking event can be calculated from Eq. 6.50 to be

$$\tilde{E}(n, s) = \frac{1}{s(1 - \tilde{f}_T(s))} = \frac{1}{s((2s)^{1/2} + \mathcal{O}(s))}. \quad (6.55)$$

The second equality is found by expanding $\tilde{f}_T(s)$ to lowest order in s . Inverting the Laplace transform of the r.h.s. of the above equation yields the asymptotic behaviour of $E(n, t)$ for large t and for $r_r < r_l$ i.e.

$$E(n, t) \approx (t/2)^{1/2}. \quad (6.56)$$

Records thus occur at a decreasing rate, as in the white noise case previously discussed. The important difference is that in the case of Brownian motion the form of the distribution of the single steps has a crucial role: When $r_r > r_l$, the walker moves to the right at the average speed $r_r - r_l$. Hence,

$$E(n, t) \approx (r_r - r_l)t. \quad (6.57)$$

6.3.1 Relation between Record and Extreme Value Statistics

From the arguments above it is clear that, even though both concerned with the largest value of some set of data, the statistics of records and that of extremes are quite different. The first deals with *when* a new record occurs, and implies a time ordering of

the data. The second focusses on the size distribution of the largest value in a set. Properties of record statistics can be obtained from those of extreme value statistics through the following observation: Consider a sequence

$$\nu_0, \nu_1, \nu_2, \nu_3, \dots, \nu_N, \nu_{N+1}, \dots$$

of independent and identically distributed stochastic variables. The probability P_N that ν_N is a record can be expressed in the following way

$$P_N = \int_{-\infty}^{\infty} \text{Prob}\{\max[\nu_0, \nu_1, \nu_2, \dots, \nu_{N-1}] < x\} P_\nu(x) dx, \quad (6.58)$$

where $P_\nu(x)$ is the probability density shared by all the ν_k 's. As we did in Eq. 6.4, we notice that

$$\text{Prob}\{\max[\nu_0, \nu_1, \nu_2, \dots, \nu_{N-1}] < x\} = [F_\nu(x)]^N, \quad (6.59)$$

where $F_\nu(x)$ of is the cumulative distribution for one of the ν variables. Furthermore, since $P_\nu(x) = dF_\nu(x)/dx$ one obtains

$$P_N = \int_{-\infty}^{\infty} [F_\nu(x)]^N \frac{dF_\nu(x)}{dx} dx = \int_{-\infty}^{\infty} F_\nu^N dF_\nu = \frac{1}{N}, \quad (6.60)$$

i.e. the same result as in Eq. 6.28. The derivation makes it explicit that although extreme value statistics depends on the nature of the probability density function of the underlying stochastic variables, the probability of a record does not. For the same reason it may be illuminating to repeat the calculation of Sec. 6.2 starting from extreme value considerations. Assume then that records occur at time steps m_i with $i = 1, 2, \dots, n-1$ in addition to the record at time step zero. We can express $P_n(l, [0, m_1, m_2, \dots, m_{n-1}])$ (see Eq. 6.33) as

$$\begin{aligned} P_n(l, [0, m_1, m_2, \dots, m_{n-1}]) &= \int_{-\infty}^{\infty} dx_{n-1} P_\nu(x_{n-1}) \int_{-\infty}^{x_{n-1}} dx_{n-2} P_\nu(x_{n-2}) \\ &\quad \int_{-\infty}^{x_{n-2}} dx_{n-3} P_\nu(x_{n-3}) \cdots \int_{-\infty}^{x_{m_2}} dx_{m_1} P_\nu(x_{m_1}) \int_{-\infty}^{x_{m_1}} dx_0 P_\nu(x_0) \\ &\quad \text{Prob}\{x_1 < x_0, x_2 < x_0, \dots, x_{m_1-1} < x_0\} \\ &\quad \text{Prob}\{x_{m_1+1} < x_{m_1}, x_2 < x_{m_1}, \dots, x_{m_2-1} < x_{m_1}\} \\ &\quad \cdots \\ &\quad \text{Prob}\{x_{m_{n-2}+1} < x_{m_{n-2}}, \dots, x_{m_{n-1}-1} < x_{m_{n-2}}\} \\ &\quad \text{Prob}\{x_{m_{n-1}+1} < x_{m_{n-1}}, \dots, x_l < x_{m_{n-1}}\}. \end{aligned} \quad (6.61)$$

We now introduce the distribution function, simplify the notation for the integration variables and obtain

$$\begin{aligned}
P_n(l, [0, m_1, m_2, \dots, m_{n-1}]) &= \int_{-\infty}^{\infty} dx_{n-1} P_\nu(x_{n-1}) \int_{-\infty}^{x_{n-1}} dx_{n-2} P_\nu(x_{n-2}) \\
&\quad \int_{-\infty}^{x_{n-2}} dx_{n-3} P_\nu(x_{n-3}) \cdots \int_{-\infty}^{x_{m_1}} dx_{m_1} P_\nu(x_{m_1}) \int_{-\infty}^{x_{m_1}} dx_0 P_\nu(x_0) \\
&\quad F_\nu(x_{m_{n-1}})^{l-m_{n-1}} F_\nu(x_{m_{n-2}})^{m_{n-1}-m_{n-2}-1} F_\nu(x_{m_{n-3}})^{m_{n-2}-m_{n-3}-1} \cdots \\
&\quad \cdots F_\nu(x_{m_1})^{m_2-m_1-1} F_\nu(x_0)^{m_1-1}.
\end{aligned} \tag{6.62}$$

Making the substitution $z_i = F_\nu(x_i)$, which we can do since $P_\nu(x_i) = dF_\nu/dx_i$, the above expression reduces to the integral

$$\begin{aligned}
P_n(l, [0, m_1, m_2, \dots, m_{n-1}]) &= \int_0^1 dz_{n-1} \int_0^{z_{n-1}} dz_{n-2} \cdots \int_0^{z_2} dz_1 \int_0^{z_1} dz_0 \\
&\quad z_{n-1}^{l-m_{n-1}} z_{n-2}^{m_{n-1}-m_{n-2}-1} z_{n-3}^{m_{n-2}-m_{n-3}-1} \cdots z_{m_1}^{m_2-m_1-1} z_0^{m_1-1}.
\end{aligned} \tag{6.63}$$

The derivation shows why $P_n(l, [0, m_1, m_2, \dots, m_{n-1}])$ is independent of the distribution of the signal from which the records are extracted. The integrals can easily be performed by starting with the integral over z_0 , followed by integration over z_1 , and so forth. The result is

$$P_n(l, [0, m_1, m_2, \dots, m_{n-1}]) = \frac{1}{l m_{n-1} m_{n-2} \cdots m_1 m_0}, \tag{6.64}$$

which is in agreement with Eq. 6.33.

6.4 Summary

In summary, the statistics of extreme values is on the one hand more general and on the other more specific than that of record values. Firstly, it is not concerned with the order in which the extremes occur and, for this reason, record statistics can be derived from it. Secondly, it computes the probability distribution of the values of the extremes, whereas record statistics is indifferent to it. Being concerned with the actual values of the extremes, extreme value statistics is sensitive to the distribution of the underlying stochastic variables. This leads to three different universality classes for the distribution of extrema. In contrast, record statistics is essentially described by the log-Poisson process irrespective of the distribution.

6.5 Tangent map intermittency

This chapter builds on the paper A. Diaz-Ruelas, H.J. Jensen, D. Piovani and A. Rolbedo, *Tangent map intermittency as an approximate analysis of intermittency in a high dimensional fully stochastic dynamical system: The Tangled Nature model*, Chaos, 26, 123105 (2016).

Intermittent dynamics in the form of long periods of little change separated by relatively short time intervals of hectic activity is observed in many complex systems. Examples include snoring, mass extinctions, financial crashes and brain activity. Such systems contain large numbers of components and very often involve stochastic processes. It was more than 40 years ago suggested by Procaccia and Schuster that important aspects of such intermittent dynamics can be captured by a simple essentially deterministic equation. It is not in general straight forward to connect in detail the complex system to the mathematics considered by Procaccia and Schuster. We consider a multi-component model of evolutionary ecology and derive how the single component equation of the type considered by Procaccia and Schuster is related to the parameters of the stochastic many component dynamics. The single component description enables us to describe aspects of the intermittent extinction dynamics that so far has eluded mathematical analysis. We also demonstrate that the single component mathematics is able to qualitatively mimic the evolution and extinction dynamics of consecutive ecologies generated by the full many component model. We think that our results expands the applicability of the analysis put forward by Procaccia and Schuster and thereby help connect the methodology developed for deterministic and typically low dimensional dynamics to the stochastic dynamics of complex systems.

High-dimensional complex systems, such as turbulence, relaxing glasses, biological evolution, the financial market or brain dynamics, exhibit intermittent dynamics [67, 65]. While intermittency in basic one-dimensional non-linear maps at the so-called tangent bifurcation [65] has received significant attention e.g. because of their universal aspects [33] and has been suggested as a universal mechanism for $1/f$ noise [56]. In fact the paradigmatic Pomeau-Manneville map [54] was derived to represent intermittency in weakly turbulent fluid dynamics. The relevance of such maps to high-dimensional stochastic systems depends on whether a robust macroscopic degree of freedom emerges, which is able to capture the dominant dynamics.

A case in point is the Tangled Nature (TaNa) model [67] of evolutionary ecology, since it displays intermittent evolution at the macroscopic level while microscopically individuals reproduce, mutate and die at essentially constant rates [14, 30]. Numerical simulations of the model show that the total population $N(t)$ as a function of time t (in the scale of generations) consists of quasi-stable, steady, periods that alternate with interludes of hectic transitions, during which $N(t)$ exhibits large amplitude fluctua-

tions [14, 30]. The populations of species behave accordingly, during the quasi-stable periods they predominantly retain their identity, but at the transitions some species vanish, others arise, while the rest survive [14, 30].

Here we study the incidence of intermittency, as displayed close to the tangent bifurcation in low-dimensional nonlinear maps, in the macroscopic behavior of the TaNa model. We make two intents. The first one is to approximate the evolution equations of the model, via determination of mean-field lowest-order local terms to obtain a map near tangency that reproduces the prototypical quasi-stable episode. The second is to model, phenomenologically, the sequences of consecutive quasi-stable and hectic periods via a nonlinear dynamical model that makes use of the families of tangent bifurcations that occur in one-dimensional quadratic maps. We arrive at the following general picture. The dynamics of the original TaNa model is fully stochastic and fluctuations are very important in steering its progress. The one-dimensional mean-field local map near tangency we derive, is deterministic and has the limitation, without further development, that only one quasi-stable event can be generated at a time. But this map offers criteria for the duration of the quasi-stable episodes in terms of the TaNa model parameters. Based on these criteria we introduce a stochastic element in the construction of an otherwise deterministic nonlinear dynamical model. This model reproduces qualitatively the sequences of quasi-stable periods with varying types and numbers of species as in the TaNa model. We comment on the class of nonlinear dynamical systems with this kind of stochastic element. We reach the conclusion that in spite of the approximations incurred and assumptions made our study facilitates interesting insights that hint to a radical reduction of degrees of freedom under certain circumstances.

6.6 The Tangled Nature model

The Tangled Nature model is a model of evolutionary ecology, which studies the macro-dynamics emerging from the dynamics of individual organisms or agents, co-evolving together and subject to a web of mutual interactions. The model is an attempt to identify possible simple mechanisms behind the myriad of complicated interactions, feedback loops, contingencies, etc., as one moves from the short time reproductive dynamics at the level of individuals, to the long time systems level behaviour. The strategy is to keep the model sufficiently simple to enable analysis, and to pinpoint the details or assumptions in the model that are responsible for the specific behaviour at the systems level. One major concern of the model has been to understand how the smooth continuous pace of the reproductive dynamics at the level of individuals, can lead to intermittent or punctuated dynamics at the level of high taxonomic structures.

The model was introduced in [14, 30] and since then, the model framework has been used by several authors see *e.g.* [61, 79, 7, 51, 75]. A summary of some of the models features and predictions can be found in [67].

Description of the model

The dynamical entities of the TaNa model consist of agents represented by a sequence of binary variables with fixed length L [32]. We denote by $n(\mathbf{S}^a, t)$ the number of agents of type $\mathbf{S}^a = (S_1^a, S_2^a, \dots, S_L^a)$ (here $S_i^a \in \{-1, 1\}$) at time t and the total population is $N(t) = \sum_{a=1}^{2^L} n(\mathbf{S}^a, t)$. A time step is defined as a succession of one annihilation and of one reproduction attempt. Annihilation consists of choosing an agent at random with uniform probability and remove the agent with probability p_{kill} , taken to be constant in time and independent on the type. Reproduction: choose with uniform probability an agent, \mathbf{S}^a , at random and duplicate the agent (and remove the mother) with probability

$$p_{off}(\mathbf{S}^a, t) = \frac{\exp(H(\mathbf{S}^a, t))}{1 + \exp(H(\mathbf{S}^a, t))}, \quad (6.65)$$

which depends on the occupancy distribution of all the types at time t through the weight function

$$H(\mathbf{S}^a, t) = \frac{k}{N(t)} \sum_b J(\mathbf{S}^a, \mathbf{S}^b) n(\mathbf{S}^b, t) - \mu N(t). \quad (6.66)$$

In Eq. (6.66), the first term couples the agent \mathbf{S}^a to one of type \mathbf{S}^b by introducing the interaction strength $J(\mathbf{S}^a, \mathbf{S}^b)$, whose values are randomly distributed in the interval $[-1, +1]$. For simplification and to emphasize interactions we here assume: $J(\mathbf{S}^a, \mathbf{S}^a) = 0$. The parameter k scales the interactions strength and μ can be thought of as the carrying capacity of the environment. An increase (decrease) in μ corresponds to harsher (more favourable) external conditions.

Mutations occur in the following way: For each of the two copies \mathbf{S}^{a_1} and \mathbf{S}^{a_2} , a single mutation changes the sign of one of the genes: $S_i^{a_1} \rightarrow -S_i^{a_1}$, $S_i^{a_2} \rightarrow -S_i^{a_2}$ with probability p_{mut} . We define a generation to consist of $N(t)/p_{kill}$ time steps, *i.e.* the average time needed to kill all the individuals at time t . These microscopic rules generate intermittent macro dynamics[30] as shown in Fig. 6.1. The long quiescent epochs are called quasi Evolutionary Stable Strategies (qESS), since they do remind one of John Maynard Smith's notion of Evolutionary Stable Strategies introduced in his game theoretic description of evolution [69].

The weight function H will fluctuate about the value given by the stable dynamical fixed point condition $p_{off}(H) = p_{kill}$. This suggests that the mean field value of H

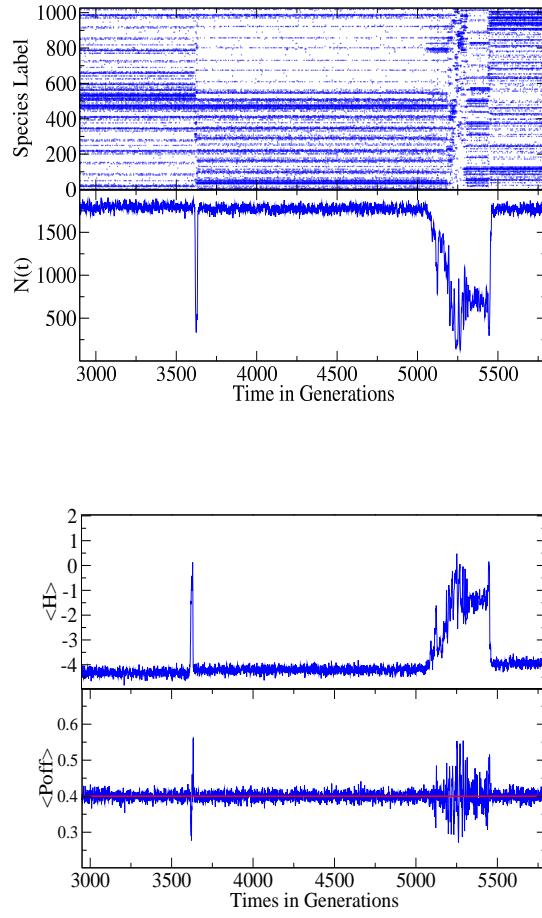


Figure 6.1: Left Panel: Total population as a function of time (in generations) for a single realization of the TaNa model. The punctuated dynamics is clearly visible: quasi-stable periods alternate with periods of hectic transitions, during which $N(t)$ exhibits large amplitude fluctuations. Right panel: The average of the weight function H and the reproduction probability. The parameters are $L = 10$, $p_{kill} = 0.4$, $p_{mut} = 0.02$, $\mu = 0.007$, $k = 40$ the red line indicates p_{kill} .

may indeed evolve in an intermittent way that may be captured by a tangent map. We will therefore derive the mean field map for $\langle H \rangle$.

Derivation of Mean Field map for H

To establish a map for the mean field approximation to the weight function H , we need to analyse each of the microscopic stochastic processes that can lead to a change in H . These are reproduction, with or without mutation and death. And we will make use of the fact that if a quantity, say X , undergoes the change to $X \mapsto X + \Delta$ with probability p and remains unchanged $X \mapsto X$ with probability $1 - p$, then in the mean field approximation we have $\langle X \rangle \mapsto \langle X \rangle + p\Delta$. We use a short hand notation in which we label individuals and types as $i, j, k \dots$ and accordingly the interaction between two types i and j as J_{ij} .

Reproduction with no mutation. We need to estimate the average change to the likelihood function, H_i for type i given that an individual of type j_0 reproduces without mutating. The change in H_i is given by

$$\begin{aligned} H_i &\mapsto \frac{k}{N+1} \left[\sum_{j \neq j_0} J_{i,j} n_j + J_{i,j_0} (n_{j_0} + 1) \right] - \mu(N+1) \\ &= \frac{k}{N+1} \left(\sum_j J_{i,j} n_j - \mu N \right) + \left(\frac{k}{N+1} J_{i,j_0} - \mu \right) \\ &= H_i + \Delta_{R,0m}^i(j_0) \end{aligned} \quad (6.67)$$

We replaced $N + 1$ by N in the first term and introduced the change

$$\Delta_{R,0m}^i(j_0) = \frac{k}{N+1} J_{i,j_0} - \mu, \quad (6.68)$$

which will occur with probability

$$p_{R,0m}^i(j_0) = \frac{n_{j_0}}{N} p_{off}(j_0) (P_{mut}^{(0)})^2, \quad (6.69)$$

where $P_{mut}^{(0)} = (1 - p_{mut})^L$ is the probability of no mutations occurring, and its counted twice, once for each offspring. Averaging over all possible types (of which there are $\Omega = 2^L$) we obtain

$$\begin{aligned} \bar{\Delta}_{R,0m} &= \langle \Delta_{R,0m}^i(j_0) \rangle = \frac{1}{\Omega} \sum_{j_0} \left(\frac{k}{N+1} J_{i,j_0} - \mu \right) \frac{n_{j_0}}{N} p_{off}(j_0) (P_{mut}^{(0)})^2 \\ &\mapsto \left(\frac{k \bar{J}}{N+1} - \mu \right) \langle p_{off} \rangle_{ext} (1 - p_{mut})^{2L}, \end{aligned} \quad (6.70)$$

where we have introduced \bar{J} , which denotes the strengths $J_{i,j}$ averaged over pairs of interacting extant types and similarly $\langle p_{off} \rangle_{ext}$ denotes the offspring probability average over extant types.

Reproduction with 1 mutation Next we consider the average change to the likelihood function, H_i for type i given that an individual of type j_0 reproduces with one copy mutating and ending in q_o and the other not mutating. The change in H_i is given by

$$\begin{aligned} H_i &\mapsto \frac{k}{N+1} \left[\sum_{j \neq q_0} J_{i,j} n_j + J_{i,q_0} (n_{q_0} + 1) \right] - \mu(N+1) \\ &= \frac{k}{N+1} \left(\sum_j J_{i,j} n_j - \mu N \right) + \left(\frac{k}{N+1} J_{i,q_0} - \mu \right) \\ &= H_i + \Delta_{R,m}^i(q_0). \end{aligned} \quad (6.71)$$

Again we have replaced $N+1$ by N in the first term and introduced the change

$$\Delta_{R,m}^i(q_0) = \frac{k}{N+1} J_{i,q_0} - \mu, \quad (6.72)$$

which will occur with probability

$$p_{R,m}^i(j_0) = \frac{n_{j_0}}{N} p_{off}(j_0) p_{j_0 \rightarrow q_0}, \quad (6.73)$$

where

$$p_{j_0 \rightarrow q_0} = 2p_{mut}^{d_{j_0 q_0}} (1 - p_{mut})^{L - d_{j_0 q_0}}, \quad (6.74)$$

and $d_{j_0 q_0}$ is the hamming distance between the sequences j_0 and q_0 . This means that

$$\bar{\Delta}_{R,1m}^i = \sum_{j \neq q_0} \left(\frac{k}{N+1} J_{i,j} - \mu \right) \frac{n_{j_0}}{N} p_j^{off} p_{mut}^{d_{j_0 q_0}} (1 - p_{mut})^{L - d_{j_0 q_0}}. \quad (6.75)$$

By limiting our approximation to the nearest neighbours, and proceeding like in the previous case, we obtain

$$\bar{\Delta}_{R,1m}^i = L p_{mut}^{(o)} p_{mut}^{(1)} \left(\frac{k \tilde{J}}{N} - \mu \right) \langle p_{off} \rangle_{ext}, \quad (6.76)$$

where L is the number of first neighbours and $P_{mut}^{(1)} = p_{mut} (1 - p_{mut})^{(L-1)}$ denotes the probability that exactly one L genes mutate. Notice the difference between \bar{J} introduced in Eq. (6.70) and the averaged quantity \tilde{J} introduced in this equation. The

two differs by being averages over different sets of types. Here \tilde{J} is averaged over interaction strengths J_{ij} connecting already occupied type and types hit by a new mutation, *i.e.* types located in the perimeter of the cluster of extant reproducing sites. In contrast \bar{J} is the average of the interaction strength between extant types. We will expect that typically $\tilde{J} < \bar{J}$ because adaptation has favoured mutualistic interactions amongst the extant types. However, an accurate estimate of the two quantities from first principle is of course very difficult.

Reproduction with 2 mutations. Next we consider the average change to the likelihood function, H_i for type i given that an individual of type j_0 reproduces with both copies mutating and ending in q_0 and q_1 . The change in H_i is given by

$$\begin{aligned} H_i &\mapsto \frac{k}{N+1} \left[\sum_{j \neq q_0, q_1, j_o} J_{i,j} n_j + J_{i,q_0}(n_{q_0} + 1) + J_{i,q_1}(n_{q_1} + 1) + J_{i,q_1}(n_{q_1} - 1) \right] - \mu(N+1) \\ &= \frac{k}{N+1} \left(\sum_j J_{i,j} n_j - \mu N \right) + \left(\frac{k}{N+1} (J_{i,q_0} + J_{i,q_1} - J_{i,j_o}) - \mu \right) \\ &= H_i + \Delta_{R,2m}^i(q_0), \end{aligned} \tag{6.77}$$

where we have consider the fact that the number of individuals of the parent decreases in case of 2 mutations. Again we have replaced $N+1$ by N in the first term and introduced the change

$$\Delta_{R,2m}^i(q_0) = \frac{k}{N+1} (J_{i,q_0} + J_{i,q_1} - J_{i,j_o}) - \mu, \tag{6.78}$$

which will occur with probability

$$p_{R,m}^i(j_0) = \frac{n_{j_0}}{N} p_{off}(j_0) p_{j_o \rightarrow q_0} p_{j_o \rightarrow q_1}. \tag{6.79}$$

And once again limiting our approximation to the nearest neighbour mutations we obtain

$$\bar{\Delta}_{R,2m} = L^2 (P_{mut}^{(1)})^2 \left(\frac{k \tilde{J}}{N} - \mu \right) \langle p_{off} \rangle_{ext}. \tag{6.80}$$

Notice the difference between \bar{J} introduced in Eq. (6.70) and the averaged quantity \tilde{J} introduced in this equation. The two differs by being averages over different

sets of types. Here \tilde{J} is averaged over interaction strengths J_{ij} connecting already occupied type and types hit by a new mutation, *i.e.* types located in the perimeter of the cluster of extant reproducing sites. In contrast \bar{J} is the average of the interaction strength between extant types. We will expect that typically $\tilde{J} < \bar{J}$ because adaptation has favoured mutualistic interactions amongst the extant types. This is in fact verified by our simulations, see Fig. 6.3 below.

Killing event on site j_0 leads to

$$\begin{aligned} H_i &\mapsto \frac{k}{N-1} \left[\sum_{j \neq j_0} J_{i,j} n_j + J_{i,0_0}(n_{j_0} - 1) \right] - \mu(N-1) \\ &= \frac{k}{N-1} \left(\sum_j J_{i,j} n_j - \mu N \right) - \left(\frac{k}{N} J_{i,j_0} - \mu \right) \\ &= H_i - \Delta_{R,0m}^i(j_0). \end{aligned} \quad (6.81)$$

This change occurs with probability $(n_{j_0}/N)p_{kill}$.

Combining this result with the weighted results in Eqs. (6.70), (6.76) and (6.80) we obtain the following map, which in mean field describes how $\langle H \rangle$ changes as an effect of the microscopic reproduction and killing events

$$\langle H \rangle \mapsto \langle H \rangle + A \langle p_{off} \rangle_{ext} - B p_{kill}, \quad (6.82)$$

where the coefficients are given by

$$A = \left(\frac{k\bar{J}}{N} - \mu \right) (P_{mut}^{(0)})^2 + \left(\frac{k\tilde{J}}{N} - \mu \right) (P_{mut}^{(0)} + Lp_{mut}^{(1)}) Lp_{mut}^{(1)} \quad (6.83)$$

$$B = \frac{k\bar{J}}{N} - \mu \quad (6.84)$$

We have derived a map for the evolution of $\langle H \rangle$. We now need to close the map, *i.e.* we need a way to express the H^i dependency of $\langle p_{off} \rangle_{ext}$ in terms of $\langle H \rangle$. We could assume

$$\langle p_{off}(H^i) \rangle_{ext} \mapsto p_{off}(\langle H \rangle_{ext}). \quad (6.85)$$

This procedure gives us the following map for x_n (which we use as shorthand for the iterates of $\langle H \rangle_{ext}$)

$$x_{n+1} = x_n + A p_{off}(x_n) - B p_{kill}. \quad (6.86)$$

If $0 < \frac{B}{A}p_{kill} < 1$ the map in Eq. (6.86) will have a fixed point x^* given by $p_{off}(x^*) = Bp_{kill}/A$. The fixed point is stable if $A > 0$ and $B > 0$ and attractive if $A < 0$ and $B < 0$. So for $0 < \frac{B}{A}p_{kill} < 1$ this map will not lead to tangent intermittency. For $\frac{B}{A}p_{kill} < 0$ or $\frac{B}{A}p_{kill} > 1$ x_n depends monotonously on n and the scenario needed for tangent intermittency could exist. However, values extracted from the numerical analysis suggest that $0 < \frac{B}{A}p_{kill} < 1$.

Therefore we conclude that the dramatic mean field approximation suggested in Eq. (6.85), which corresponds to the replacement $\langle H^n \rangle_{ext} \mapsto \langle H \rangle_{ext}^n$ for all $n \in \mathbb{N}$, wipes out the intermittency. To establish a mean field description of the intermittency we instead expand $p_{off}(H^i)$ in Eq. (6.82) to second order about x^* and replaces only $\langle H^2 \rangle_{ext}$ by $\langle H \rangle_{ext}^2$. This leads to a tangent map and we study the intermittency of this map in the next section.

6.7 Analysis of the map in the neighbourhood of tangency

We expand $p_{off}(H)$ in Eq. (6.82) to second order about $H^* = \ln[p_{kill}/(1 - p_{kill})]$,

$$p_{off}(H) = a_0 + a_1(H - H^*) + a_2(H - H^*)^2 \quad (6.87)$$

where

$$\begin{aligned} a_0 &= p_{kill}, \\ a_1 &= p'_{off}(H^*) = p_{kill}(1 - p_{kill}), \\ a_2 &= \frac{1}{2}p''_{off}(H^*) = \frac{1}{2}a_1(1 - 2p_{kill}). \end{aligned}$$

We substitute Eq. (6.87) into Eq. (6.86) and obtain the following map for $\Delta = \langle H \rangle - H^*$

$$\Delta_{n+1} = b_0 + b_1\Delta_n + b_2\Delta_n^2 \equiv f(\Delta_n), \quad (6.88)$$

where

$$\begin{aligned} b_0 &= a_0(A - B), \\ b_1 &= 1 + a_1A, \\ b_2 &= a_2A. \end{aligned}$$

Let Δ_c be given by $f'(\Delta_c) = 1$ and $\epsilon = f(\Delta_c) - \Delta_c$, i.e. at Δ_c the map has a tangent parallel to the identity and the vertical distance to the identity at this point is ϵ and is

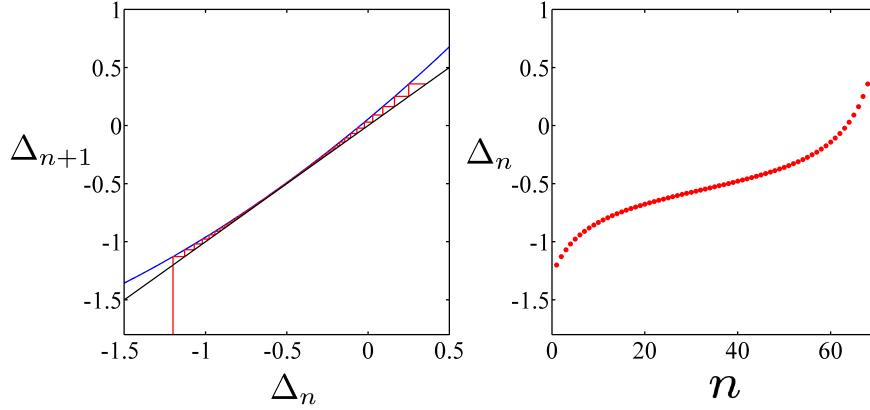


Figure 6.2: The left panel shows the first 67 iterations of the map in Eq. (6.88), with initial condition $\Delta_0 = -1.2$. The corresponding trajectory is shown in the right panel. The set of parameter values is the same than that as for Fig. 6.1, with the corresponding averaged interactions $\bar{J} \approx 0.0587$ and $\tilde{J} \approx -0.000001$, thus yielding the coefficients of the map $b_0 \approx 0.060784$, $b_1 \approx 1.167990$ and $b_2 \approx 0.151191$.

given by

$$\epsilon = b_0 - \frac{(1 - b_1)^2}{4b_2}. \quad (6.89)$$

In Figure 6.2 we show an example of an iteration of the map in Eq. (6.88) for a set of typical simulation parameters.

The number of iterations T needed to pass through the bottleneck between the map and the identity is of order $T = \pi/\sqrt{\epsilon b_2}$ (see e.g. [65] Chap. 5). Hence we have

$$\left(\frac{\pi}{T}\right)^2 = b_0 b_2 - \frac{1}{4}(1 - b_1)^2 \quad (6.90)$$

We can simplify this expression by only working to the lowest order in the killing probability and further more we will only include mutation processes considered above, *i.e.* single gene mutations in one or in both offspring. Let us denote by P_0 the probability that no mutation occur, *i.e.* $P_0 = (1 - p_{mut})^{2L}$. Since we neglect all other mutation events than the two kinds just described, we have the approximation

$$1 - P_0 = (P_{mut}^{(0)} + Lp_{mut}^{(1)})Lp_{mut}^{(1)}, \quad (6.91)$$

in which case Eq. (6.83) becomes

$$A = \left(\frac{k\bar{J}}{N} - \mu \right) P_0 + \left(\frac{k\tilde{J}}{N} - \mu \right) (1 - P_0). \quad (6.92)$$

With these approximations we arrive at

$$\left(\frac{\pi}{T} \right)^2 \simeq -\frac{k}{2N} (\bar{J} - \tilde{J})(1 - P_0) \left[\frac{k}{N} (\bar{J} - \tilde{J})P_0 + \frac{k}{N} \tilde{J} - \mu \right] p_{kill}^2. \quad (6.93)$$

We find that $kP_0(\bar{J} - \tilde{J})/N$ is very small and hence that the expression for $(\pi/T)^2$ is well approximated by

$$\left(\frac{\pi}{T} \right)^2 \simeq -\frac{k}{2N} (\bar{J} - \tilde{J})(1 - P_0) \left(\frac{k}{N} \tilde{J} - \mu \right) p_{kill}^2. \quad (6.94)$$

We conclude that our mean field analysis suggests that the length of the qESS, *i.e.* the metastable quiescent epochs, is set by four mechanisms. Firstly, it is obvious that the rate of mutations $(1 - P_0)$ influences the duration of the qESS: no mutations leads to no transitions and hence lead to $T = \infty$. Secondly, the rate of killing, *i.e.* the factor p_{kill}^2 . The rate of killing is related to the rate of offspring production, since due to the environmental coupling term $-\mu N$ in Eq. (6.66) on average there is a balance between reproduction and killing. Hence if p_{kill} decreases the rate of reproduction decreases and fewer mutations are produced leaving the qESS more stable and T larger. The third mechanism influencing T is the mismatch between the characteristic interaction strength between the set of extant types and mutant types located in the perimeter of the set of extant types, *i.e.* the factor $(\bar{J} - \tilde{J})$. One may consider this factor as being related to the selective pressure on the qESS state. Namely, if $\bar{J} = \tilde{J}$ there is no selective pressure, since the mutants are entirely like the wild types. The lack of selective pressure drives T to become infinite. Finally the term $k\tilde{J}/N - \mu$. Again one may see the fact that when $k\tilde{J}/N = \mu$ the duration T becomes infinite as representing the fact that a mutant population for which $\tilde{J}/N = \mu$ is already well tuned to the environmental pressure represented by μ .

It is of course interesting to try to relate the prediction for the duration T given by Eq. (6.94) to the actual qESS intermittency observed in simulations of the Tangled Nature model. To do this we have determined the time average values of \bar{J} and \tilde{J} during 4000 qESS states and plot the distributions obtained in Fig. 6.3. In order for Eq. (6.94) to make sense the sign has of course to be positive. The two terms that could in principle be problematic are $(\bar{J} - \tilde{J})$ and $(\frac{k\tilde{J}}{N} - \mu)$. In Fig. 6.3 we show the

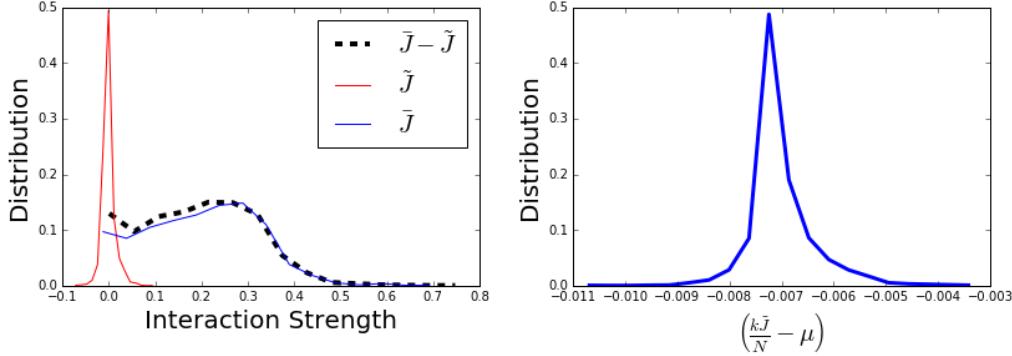


Figure 6.3: The distribution of factors that enter into the expression Eq. (6.94) for the duration of the qESS. The \bar{J} and \tilde{J} are the time averaged during a qESS. The data show is sampled over 4000 qESS realisations. The parameters for the simulation is the same as in Fig. 6.1.

distribution of these terms obtained from a set of 4000 measured qESS. The difference in the average J 's is always positive while the second term is alway negative. Given the high stochasticity of the TaNa, this doesn't mean that a specific qESS cannot produce a negative value, but that this is extremely unlikely, and perhaps it wouldn't last long enough for us to notice it. Amongst all the realisations of \bar{J} and \tilde{J} we recorded only 0.16% corresponds to a negative right hand side in Eq. (6.94).

Indeed one has to consider |Eq.(30) as a qualitative description of the duration more than a quantitative one. This is the case because of the approximations we were forced to make when deriving the mean field expressions. Moreover, it is difficult to relate the quantitative values for T obtained from Eq. (30) to the durations observed in the simulations because of the always difficult problem of relating mean field time to Monte Carlo simulation time steps.

Chapter 7

Percolation

The process consisting of percolation on a lattice is another paradigmatic model, which conceptually is of great relevance to complex systems[71]. Consider a d dimensional hyper-cubic lattice of size $N = L^d$. Assume that a site is occupied (by whatever you fancy: a coin, an elephant, a tree, ...) with probability p . On average the number of occupied sites is $N_{occ} = pN$. We define cluster as sets of occupied sites that are connected though nearest neighbour adjacency. See Fig. 7.1.

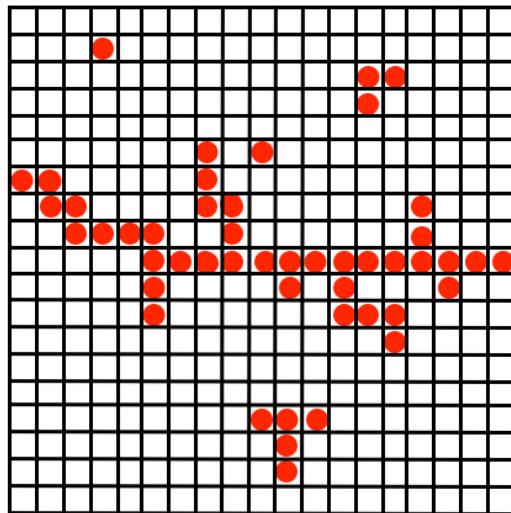


Figure 7.1: Cluster of occupied sites. We see two cluster of size 1, one of size 3, one of size 5 and one systems spanning cluster and

This is a purely geometric process. No interaction is assume between the sites, so

in this respect the situation is simple than, say, when dealing with the Ising model. We will discuss the distribution of cluster sizes[13]. The number os clusters of size s per lattice site is given by

$$n_s(p) = \frac{\# \text{ clusters of size } s}{N} \quad (7.1)$$

The probability that a random site belongs to a cluster of size s is then given by $sn_s(p)$, which is normalised in the following way

$$\sum_{s=1}^{\infty} sn_s(p) = \frac{\# \text{ occupied sites}}{N} = \frac{N_{occ}}{N} = p. \quad (7.2)$$

The total number of clusters, N_{cl} , is given by

$$N_{cl}(p) = \sum_{s=1}^{\infty} n_s(p). \quad (7.3)$$

The cluster size distribution $n_s(p)$ turns out behave like

$$n_s(p) \propto s^{-\tau} \exp(-s/s_0) \quad (7.4)$$

for values of s large compared to 1. Like in the Ising model percolation also involves a phase transition. In this case the transition is between low a region of low values of p for which no system spanning cluster exist and a region of high values of p for which a system spanning cluster does exist. These two regions are separated by a sharp phase transition (rather than a gradual change) at a critical value p_c . At p_c the average cluster size diverges as $N \rightarrow \infty$ and the probability P_∞ , that an occupied site belongs to the system spanning (or infinite) cluster becomes non-zero.

The exponent τ of the cluster size distribution $n_s(p)$ in Eq. (7.4) is $\tau = 5/2$ in mean field and the exact result in 2 dimensions is $\tau = 187/91$. As the critical value p_c is approached the characteristic cluster size s_0 diverges as a power law, which in mean field is given by $s_0 \propto |p - p_c|^{-2}$.

A simple one dimensional analysis suggests how the origin of exponential factor in Eq. (7.4). Namely think of a one dimensional lattice. The probability that a randomly selected site belongs to a cluster of size s will be given by

$$sn_s(p) = s(1-p)p\ldots pp(1-p) = s(1-p)^2 p^s. \quad (7.5)$$

Here the first factor s is the number of ways a t random to land on a cluster of size s and the two factors $(1-p)$ correspond to the vacant sites at either end of the s consecutively occupied sites, which happens with probability p^s . So we arrive a t

$$n_s(p) = (1-p)^2 \exp(-s/s_0). \quad (7.6)$$

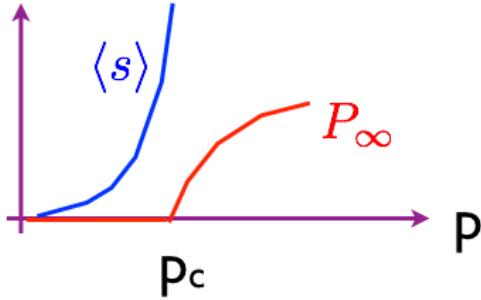


Figure 7.2: Sketch of the percolation transition at p_c .

Here the characteristic cluster size is given by $s_0 = 1/\ln(1/p)$ or for p close to 1:

$$s_0 = \frac{-1}{\ln(p)} = \frac{-1}{\ln[1 - (1-p)]} \simeq \frac{1}{1-p}, \quad (7.7)$$

and we have an algebraic divergence as one approach the critical value $p_c = 1$. It is clear that in one dimension the percolation transition clearly takes place at the someone trivial value $p = 1$ where all sites are occupied, since for $p < 1$ we cannot have a system spanning cluster.

In Fig. 7.3 we show a figure from the very informative book by Christensen and Moloney *Complexity and Criticality* [13]. One notice how a cluster able to span the entire system appears at p_c .

One can easily perform a mean field analysis of percolation by the following consideration. The probability that an occupied site *does not* belong to the spanning cluster is given by $1 - P_\infty$. The only way a site can avoid not to be part of the spanning cluster is if none of its nearest neighbour sites belongs to the spanning cluster. The probability for all q nearest neighbour sites not to belong to the spanning cluster is, when we neglect correlations, given by $(1 - pP_\infty)^q$. Namely pP_∞ is the probability that the site is occupied and belongs to the spanning cluster. In this mean field approximation P_∞ is the solution of the equation

$$1 - P_\infty = (1 - pP_\infty)^q := f(P_\infty). \quad (7.8)$$

It is straight forward to see graphically that for $p < p_c = 1/q$ only the solution $P_\infty = 0$ exists. And that for $p > p_c$ a non-zero solution appear. For p above but close to p_c we have P_∞ close to zero and we can approximate the right hand side of Eq. (7.8) by

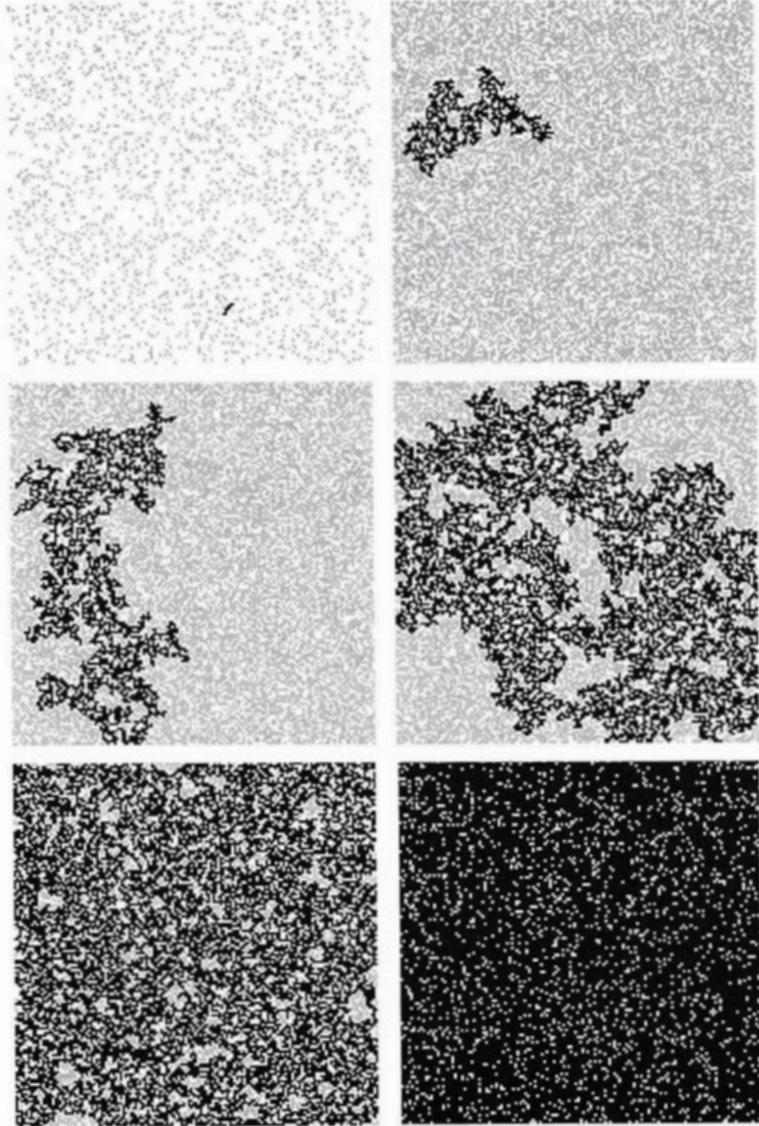


Fig. 1.2 Six realisations of two-dimensional site percolation on a square lattice of size $L = 150$ for occupation probabilities $p = 0.10, 0.55, 0.58, 0.59274621, 0.65, 0.90$, from left to right and top to bottom, respectively. Occupied sites are dark grey while unoccupied sites are white. For each realisation, the largest cluster has been shaded black.

Figure 7.3: Cluster of occupied sites on a two dimensional square lattice for different values of p . From page 4 in Christensen and Moloney *Complexity and Criticality*, Imperial College Press, 2005.

Taylor expanding $f(x)$ about $x = 0$. This leads to the solution given by

$$P_\infty \simeq \frac{2q^2}{q-1}(p - p_c). \quad (7.9)$$

Figure 7.4 contains a plot of the numerical solution of Eq. 7.8 for the number of nearest neighbours $q = 4$ corresponding to the 2d square lattice case.

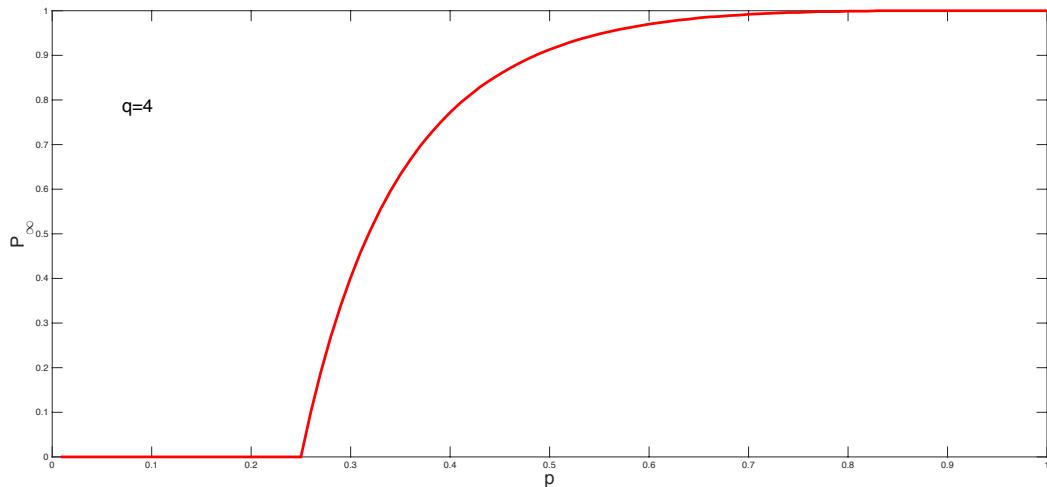


Figure 7.4: Numerical solution of P_∞ in the case $q = 4$ corresponding to 2d

We can also get a feeling for how the occupied site are correlated by means of the following measure. Let $C(\mathbf{r})$ denote the probability that if pick an occupied site at position \mathbf{r}_0 and another occupied site at position $\mathbf{r}_0 + \mathbf{r}$ they will belong to the same cluster of occupied sites.

It is straight forward to analyse $C(r)$ in one dimension. Since we assume the system is homogeneous, we can might as well choose $\mathbf{r}_0 = 0$. We then have in one dimension

$$\begin{aligned} C(r) &= \text{prob}\{x = 0, 1, \dots, r \text{ are all occupied}\} \\ &\quad + \text{prob}\{x = 0, -1, \dots, -r \text{ are all occupied}\} \\ &= 2p^r. \end{aligned} \tag{7.10}$$

This we can write as

$$C(r) = 2 \exp(r \ln(p)) = 2 \exp(-r/\xi), \tag{7.11}$$

where we introduced the correlation length

$$\xi = \frac{-1}{\ln(p)} = \frac{-1}{\ln(1 - (1-p))} \simeq \frac{1}{1-p}. \tag{7.12}$$

The last approximation is of course for p close to 1. And since $p_c = 1$ we notice that $\xi \simeq (p_c - p)^{-1}$, i.e. that the correlation length diverges algebraically with an exponent $\nu = 1$.

Chapter 8

Time correlations

In this section we describe an example of emergence in time. We first focus on very long time correlation in a stationary time signal. We will assume that interaction between the components of our system forces these to diffuse around, rather than to move around in a ballistic manner. The result is a time signal that contains very strong correlations and is characterised by what is denoted a $1/f$ power spectrum[45]. Next we discuss non-stationary time signals and briefly study the Hurst exponent.

8.1 Switching signal

We want to study correlations in a time signal $f(t)$. We measure the signal again and again at two times separated by T time units. To make life simple we will *neglect* the normalisation factor in the empirical averages, i.e. we don't divide by the number of terms in the sum in Eq. (8.1) below. A justification for this is that we are interested in the functional dependence of the correlations on the time interval T and not so much interested in the actual specific value of the correlation coefficient. Since the correlation coefficient will depend on T we talk about the correlation function. Moreover, since we are correlating the signal with itself we talk about the autocorrelation function given by:

$$\begin{aligned} C(T) &= \sum_t [f(t) - \langle f(t) \rangle][f(t+T) - \langle f(t+T) \rangle] \\ &= \int dt [f(t) - \langle f(t) \rangle][f(t+T) - \langle f(t+T) \rangle] \\ &= \int dt [f(t) - \langle f(t) \rangle][f(t+T) - \langle f(t) \rangle]. \end{aligned} \tag{8.1}$$

In the last equality we made use of the fact that the average of value $f(t)$ and $f(t + T)$ are identical.

As a way to familiarise ourselves with the autocorrelation function of a time consider a two valued signal $f(t) \in \{-A, A\}$ with $\inf t < t < \infty$. Assume that the probability that $f(t)$ switch during the time interval dt is constant and given by νdt . We want to compute the auto-correlation function

$$C(t) = \langle f(t_0)f(t_0 + t) \rangle - \langle f(t_0) \rangle \langle f(t_0 + t) \rangle.$$

where the angular brackets denotes average over t_0 .

First we note that $f(t_0)f(t_0 + t) = A^2$, if the signal has switched an even number of times during the time t . The switching of the signal is a Poisson process and the probability that no switching has occurred during the time t is

$$p_0(t) = \exp(-\nu t). \quad (8.2)$$

This one can conclude e.g. from the observation that $p_0(t + dt) = p_0(t)(1 - \nu dt)$. The probability that the signal has switched precisely n times during the time t is given by

$$p_n(t) = \frac{(\nu t)^n}{n!} \exp(-\nu t), \quad (8.3)$$

which e.g. follows from a proof by induction in n .

Now note that since $\langle f(t) \rangle = 0$ for all t ,

$$\begin{aligned} C(t) &= \langle f(0)f(t) \rangle \\ &= A \cdot \text{prob}\{f(0) = A\}[Ap_0 - Ap_1 + Ap_2 - Ap_3 + \dots] \\ &\quad - A \cdot \text{prob}\{f(0) = -A\}[-Ap_0 + Ap_1 - Ap_2 + Ap_3 + \dots] \\ &= A^2[p_0 - p_1 + p_3 - \dots]. \end{aligned} \quad (8.4)$$

By use of the above expressions for p_n one finds

$$C(t) = A^2 \exp(-2\nu|t|). \quad (8.5)$$

The correlation decay exponentially over a characteristic time scale $\tau = 1/2\nu$ set by how likely it is for the signal to switch.

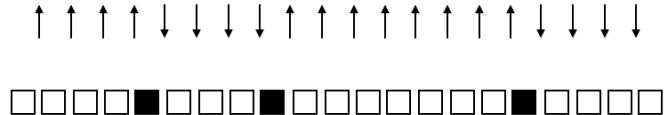
Let us briefly use this result to think about the one dimensional Ising model. We can think about spin S_i as the signal $f(t)$ with the continuous time t replaced by the integer i and $A = 1$. This signal switches when ever we encounter an antiparallel

pair of spins, see Fig. 8.1. We will consider the statistical mechanics of the chain in the following approximation corresponding to a gas of pairs of antiparallel spins. We notice that the lowest energy of an interaction bond between the spin at position i and the one at position $i + 1$ occurs when $S_i S_{i+1} = 1$ and is equal to $\Delta E_- = -J$. If the two spins have opposite sign $S_i S_{i+1} = -1$ the corresponding energy is $\Delta E_+ = J$.

At zero temperature all spins will be parallel $S_i = 1 \forall i = 1, \dots, N$ or, equally likely, $S_i = -1 \forall i = 1, \dots, N$. As the temperature is increased we can imagine that antiparallel pairs are created and that as we move along the chain antiparallel pairs appear with a probability per site ν proportional to $\exp(-\beta \Delta E_+)$, i.e.

$$\nu = \omega \exp(-\beta \Delta E_+). \quad (8.6)$$

From the similarity with the switching time signal we conclude that the spin-spin correlation function of the Ising chain will decay exponentially with spatial separation and that the correlation length will behave according to $\xi \propto 1/\nu$ and therefore diverge as $T \rightarrow 0$.



8.2 I/f signals

The autocorrelation function is an important object for the study of memory effects or causality effects in a signal. The correlation functions is equivalent to *power spectrum*. The power spectrum of the signal $f(t)$ is defined as

$$S_f(\omega) = |\hat{f}(\omega)|^2. \quad (8.7)$$

That is the absolute value squared of the Fourier transform of the signal and the power spectrum is related to the Fourier transform of the autocorrelation function:

$$S_f(\omega) = \hat{C}(\omega). \quad (8.8)$$

Here we use the following convention for the Fourier transform $\hat{g}(\omega)$ of a function $g(t)$

$$\hat{g}(\omega) = \int_{-\infty}^{\infty} dt g(t) e^{-i\omega t} \quad (8.9)$$

$$g(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \hat{g}(\omega) e^{i\omega t} \quad (8.10)$$

This relation explain why power spectra that approximately depend inversely proportional on the frequency

$$S_f(\omega) \propto 1/\omega^\beta, \quad (8.11)$$

with $\beta \simeq 1$, are of special interest [55, 77, 27]. Namely, at a somewhat heuristic level, we can substitute Eq. (8.11) into Eq. (8.8) and then into

$$C(T) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \hat{C}(\omega) e^{i\omega T}, \quad (8.12)$$

to obtain

$$C(T) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \omega^{-\beta} e^{-i\omega T} \quad (8.13)$$

$$= T^{1-\beta} \int_{-\infty}^{\infty} du u^{-\beta} e^u. \quad (8.14)$$

We made the substitution $u = \omega T$ and note that the integral in the above equation now is independent of T . So when $\beta \simeq 1$, the correlation function $C(T)$ depends very weakly on T meaning very slow decay of correlations. This indicates the particular interest in power spectra that approximately decays as one over the frequency - called $1/f$ noise. The way we carried the argument through is slightly dangerous due to possible divergent integrals; the conclusion is, however, sound.

8.2.1 Transport by Diffusion

For concreteness imagine a piece of motorway stretching from $x = -\infty$ to $x = \infty$. At $x = 0$ vehicles can enter or leave at an intersection. We will develop a model for the time evolution of the density of cars $n(x, t)$ at position x at time t . Since the cars - particles - only can leave or enter our system at $x = 0$, at all other positions, $x \neq 0$, changes during a brief time interval δ in the number of particles in a small interval $[x, x + \delta x]$ about x

$$\delta n(x, t) = n(x, t + \delta t) \delta x - n(x, t) \delta x \quad (8.15)$$

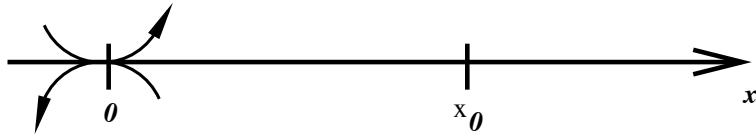


Figure 8.1: Cars/particles diffusing up and down a motorway stretching from $x = -\infty$ to $x = \infty$. At $x = 0$ an intersection allows the vehicles to enter or leave the motor way. At $x = x_0$ a traffic warden is monitoring the number of vehicles, $N(t)$, in front of him

will be caused by a difference during the time δt between the number of particles leaving the section $[x, x+\delta x]$ at $x+\delta x$ and the number of particles entering the section at x . Let $J(x, t)$ denote the particle current (number of particles crossing the position x at time t per time unit). We can then write

$$\delta n(x, t) = -J(x + \delta x, t)\delta t + J(x, t)\delta t. \quad (8.16)$$

Substituting Eq. (8.16) into Eq. (8.15) we obtain

$$n(x, t + \delta t)\delta x - n(x, t)\delta x = -J(x + \delta x, t)\delta t + J(x, t)\delta t \quad (8.17)$$

$$\Downarrow \\ \frac{\partial n(x, t)}{\partial t} = -\frac{\partial J(x, t)}{\partial x}. \quad (8.18)$$

The last equation follows in the limit $\delta x \rightarrow 0$ and $\delta t \rightarrow 0$.

This equation is exact and only assumes conservation of the particles. To obtain a closed equation for $n(x, t)$ we need to relate $J(x, t)$ to $n(x, t)$. And to do so we need to make assumptions concerning the nature of how the particles, or cars, move along the line. Let us imagine that congestion makes it impossible for cars to move freely. On the contrary assume that the drivers are forced to effectively perform random walks; i.e. diffuse along the motorway.

We emphasize that it is through this assumption that the cars, or particles, are made to interact. Here we assume that the diffusion is a result of over-crowding and interaction amongst the cars. Obviously particles might perform diffusive motion as a result of other interactions. Pollen in water diffuses because it is bombarded by large numbers of water molecules. In any case some sort of complex interaction is always responsible for diffusive motion since the particles otherwise would move around according to Newton's laws. The model might in fact be more relevant to small particles (pollen, say) suspended in a long narrow strip of water - or something else.

To model the jamming and resulting diffusive motion, we will assume that the net particle current (at coarse grained level) is from high particle density to low particle density, and linear in the density difference. We express this as

$$J(x, t) = -\gamma \frac{\partial n(x, t)}{\partial x}. \quad (8.19)$$

We combine Eq. (8.18) and Eq. (8.19) to obtain a closed form of dynamical equation for $n(x, t)$:

$$\frac{\partial n(x, t)}{\partial t} = \gamma \frac{\partial^2 n(x, t)}{\partial x^2}. \quad (8.20)$$

This is the well known diffusion equation. It describes how inhomogeneities in the density $n(x, t)$ relaxes by diffusion. The equation describes a closed system. Next we include the fact that particles might be added or removed at a certain rate $g(x, t)$ at position x at time t . According to the description above we have in particular in mind that $g(x, t)$ must describe cars leaving and entering at $x = 0$. We will later return to how this particular requirement can be imposed on $g(x, t)$. We now have our final equation of motion for $n(x, t)$

$$\frac{\partial n(x, t)}{\partial t} = \gamma \frac{\partial^2 n(x, t)}{\partial x^2} + g(x, t). \quad (8.21)$$

This is an inhomogeneous partial differential equation and we solve it easily by Fourier transformation

$$n(x, t) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \hat{n}(k, \omega) e^{i(kx + \omega t)} \quad (8.22)$$

Substitute into Eq. (8.21) to obtain an expression for $\hat{n}(k, \omega)$ in terms of the Fourier transform of the drive $\hat{g}(k, \omega)$

$$\hat{n}(k, \omega) = \frac{\hat{g}(k, \omega)}{i\omega + \gamma k^2}. \quad (8.23)$$

Now substitute Eq. (8.23) into Eq. (8.22):

$$n(x, t) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\hat{g}(k, \omega)}{i\omega + \gamma k^2} e^{i(kx + \omega t)}. \quad (8.24)$$

Next we want to focus on the density fluctuations at a specific position $x_0 > 0$. Therefore we define

$$N(t) \equiv n(x_0, t) - \langle n(x_0, t) \rangle_t, \quad (8.25)$$

where we have subtracted the temporal averaged density. We will determine the power spectrum of $N(t)$ and for this purpose need the Fourier transform

$$\begin{aligned}
\hat{N}(\omega) &= \int_{-\infty}^{\infty} dt N(t) e^{-i\omega t} \\
&= \int_{-\infty}^{\infty} dt n(x_0, t) e^{-i\omega t} - \langle n(x_0, t) \rangle_t \int_{-\infty}^{\infty} dt e^{-i\omega t} \\
&= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \int_{-\infty}^{\infty} \frac{dk}{2\pi i\omega' + \gamma k^2} \hat{g}(k, \omega') e^{ikx_0 + i\omega' t} e^{-i\omega t} - \langle n(x_0, t) \rangle_t \int_{-\infty}^{\infty} dt e^{-i\omega t} \\
&= \int_{-\infty}^{\infty} \frac{dk}{2\pi i\omega + \gamma k^2} \hat{g}(k, \omega) e^{ikx_0} - \langle n(x_0, t) \rangle_t 2\pi\delta(\omega)
\end{aligned} \tag{8.26}$$

Since we are interested in the dependence of the power spectrum for $\omega > 0$ we neglect the second term in Eq. (8.27), which, because of the delta function, will only contribute to $\hat{N}(\omega)$ at $\omega = 0$. This is how far we can go without further assumptions concerning the nature of the drive $g(x, t)$. Since this source term is meant to represent vehicles entering and leaving at position $x = 0$ we will now use

$$g(x, t) = \delta(x)\chi(t) \Rightarrow \hat{g}(k, \omega) = \hat{\chi}(\omega). \tag{8.27}$$

We then have that for $x \neq 0$ the source $g(x, t) = 0$ and at $x = 0$ the temporal variation in the flow onto and away from the “motorway” is given by $\chi(t)$. From Eq. (8.27) we get

$$\hat{N}(\omega) = \hat{\chi}(\omega) \int_{-\infty}^{\infty} \frac{dk}{2\pi i\omega + \gamma k^2} \frac{e^{ikx_0}}{i\omega + \gamma k^2} = \frac{\hat{\chi}(\omega)}{2\pi\gamma} \int_{-\infty}^{\infty} dk \frac{e^{ikx_0}}{i\frac{\omega}{\gamma} + k^2}. \tag{8.28}$$

This integral can easily be computed by contour integration in the complex k -plane. Notice that the denominator can be written as $(k - k_+)(k - k_-)$ where

$$k_{\pm} = \pm \sqrt{\frac{\omega}{\gamma}} \exp\left(\frac{3\pi}{8}i\right) = \pm \frac{1}{\sqrt{2}} \sqrt{\frac{\omega}{\gamma}} (-1 + i).$$

For $x > 0$ close the contour in the upper half plane and pick up the residue at k_+ . The power spectrum is finally calculated as the absolute value square of $\hat{N}(\omega)$

$$|\hat{N}(\omega)|^2 = \frac{|\hat{\chi}(\omega)|^2}{4\gamma\omega} e^{-\sqrt{\frac{2\omega}{\gamma}}x_0}. \tag{8.29}$$

The power spectrum of the density fluctuations is clearly influenced by the power spectrum of $\chi(t)$. Let us assume that vehicles enter and leave at the intersection in a

totally uncorrelated manner (perhaps not a totally realistic assumption) which translates into $|\hat{\chi}(\omega)|^2 = \text{constant}$. In this case

$$|\hat{N}(\omega)|^2 \propto \frac{1}{\omega} e^{-\sqrt{\frac{2\omega}{\gamma}}x_0}. \quad (8.30)$$

For frequencies so small that $\sqrt{\frac{2\omega}{\gamma}}x_0 < 1$ we have $\exp(-\sqrt{\frac{2\omega}{\gamma}}x_0) \simeq 1$ and therefore

$$|\hat{N}(\omega)|^2 \propto \frac{1}{\omega} \text{ for } \omega < \frac{\gamma}{2x_o^2} \equiv \frac{1}{2T_{diff}}. \quad (8.31)$$

Where we introduced the time scale $T_{diff} = x_o^2/\gamma$. This is the characteristic time it takes for particles, under going diffusion with a diffusion constant γ , to move from $x = 0$ to $x = x_0$.

Very long temporal correlations, as indicated by the $1/f$ behaviour of the power spectrum, is observed in very many and diverse situations: the light intensity from quasars, the ocean current, the pitch or pressure fluctuations in speech and music, the flow of traffic, the fluctuations in the resistivity of a conductor, and many more[55, 77]. Is the model we have sketched above able to explain the observed $1/f$ correlations in all these many different systems? No, probably not. Surface driven diffusion doesn't seem to be central to all these situations. The question whether a general explanation for $1/f$ exists is still an open one.

We have considered $1/f$ fluctuations here for mainly two reasons. It is a fascinating problem which is often encountered in complex systems and our discussion illustrates how one can use stochastic differential equations to go beyond equilibrium statistical mechanics to analyse temporal emergent behaviour.

To place the $1/f$ signal in perspective we return to the switching signal once again. The power spectrum is for this signal is given by the Fourier transform of the auto-correlation function in Eq. (8.5), hence

$$\tilde{C}(\omega) = \frac{2\tau A}{1 + (\tau\omega)^2}. \quad (8.32)$$

This power spectrum decays as ω^{-2} for $\omega \gg 1/\tau$ and approaches a constant as $\omega \ll 1/\tau$. This corresponds to the fact that a frequency indecent power spectrum is the Fourier transform of a δ function in time, i.e. a correlation function that decays instantaneously. At frequencies low compare to $1/2\tau$ correlations are lost, killed by the multiple switches that can have occurred for times large compare to τ .

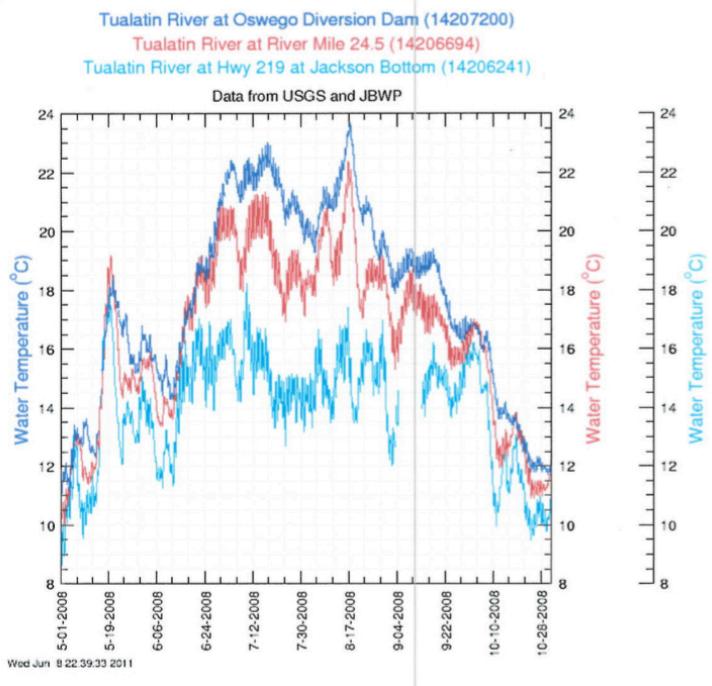


Figure 8.2: Water temperature of three rivers. Figure taken from: http://or.water.usgs.gov/grapher/tutorial/graphs/example_3site_ts.png

8.3 Non-stationary signals: Hurst exponent

The above analysis in terms of the auto-correlation function and Fourier transform assumes that the considered time series, or process, is stationary. I.e.e that the statistical properties of the process does not change with time. This is obviously often *not* the case. A simple counter example is an ordinary random walk for which, e.g, the second moment x^2 is time dependent. Time series from relay systems such as EEG recording from the brain, financial time series, temperature records etc. are examples of non-stationary time series. Figure 8.2 show examples of the temporal evolution of water temperature of three rivers.

To analyse the time correlations of such cases one needs an alternative approach. We'll briefly discuss the Hurst exponent and use a random walk with either persistence or anti-persistence between the step increments as an example.

Consider a time signal consisting of the accumulated contributions

$$X(t) = \sum_{k=1}^t s_k \quad (8.33)$$

We can think of $X(t)$ as the position of a random walker after t steps given by s_k with $k = 1, 2, \dots, t$. If the increments are independent and identically distributed (iid) described by the probability density $P(s)$ with

$$\begin{aligned} \langle s \rangle &= \sum_s s P(s), \\ \langle (\Delta s)^2 \rangle &= \sum_s (s - \langle s \rangle)^2 P(s). \end{aligned} \quad (8.34)$$

From the central limit theorem we have

$$\begin{aligned} \langle X(t) \rangle &= \langle s \rangle t, \\ \langle (\Delta X)^2(t) \rangle &= \langle (\Delta s)^2 \rangle t. \end{aligned} \quad (8.35)$$

The time dependence for signals of the form in Eq. (8.33) can be very different if the increments are not iid. The Hurst exponent is now defined as

$$\frac{\langle (\Delta X)^2(t) \rangle}{\langle (\Delta s)^2 \rangle} \propto t^{2H}. \quad (8.36)$$

So for an iid random walk with stationary increment distribution $H = 1/2$. For correlated time signals the Hurst exponent may differ from $1/2$. Figure 8.3 show examples of time series characterised by $H = 0.9, 1/2$ and 0.1 . The figure is taken from (page 182) the very readable book by J Feder, *Fractals*, Plenum Press, 1988. Values of $H > 1/2$ corresponds to signals which posses some degree of persistence where as $H < 1/2$ may be related to anti-persistence. We'll illustrate this by the following a simple random walk example below.

Originally Hurst suggested the exponent H as a way to characterise long time behaviour [20] and devised the so-called Rescaled Range analysis to determine H as defined by Eq. (8.36). Alternatively people often use a definition more applicable to short time[29]. Namely

$$\langle (X(t) - X(t_0))^2 \rangle \propto (t - t_0)^{2H}. \quad (8.37)$$

If the correlations in the process are homogeneous from short to long time the two definitions are equivalent. The second definition is often used as a phenomenological way to discuss changes in the correlations of a time signal in terms of a time dependent Hurst exponent computed by means of Eq. (8.37) for short time differences $t - t_0$.

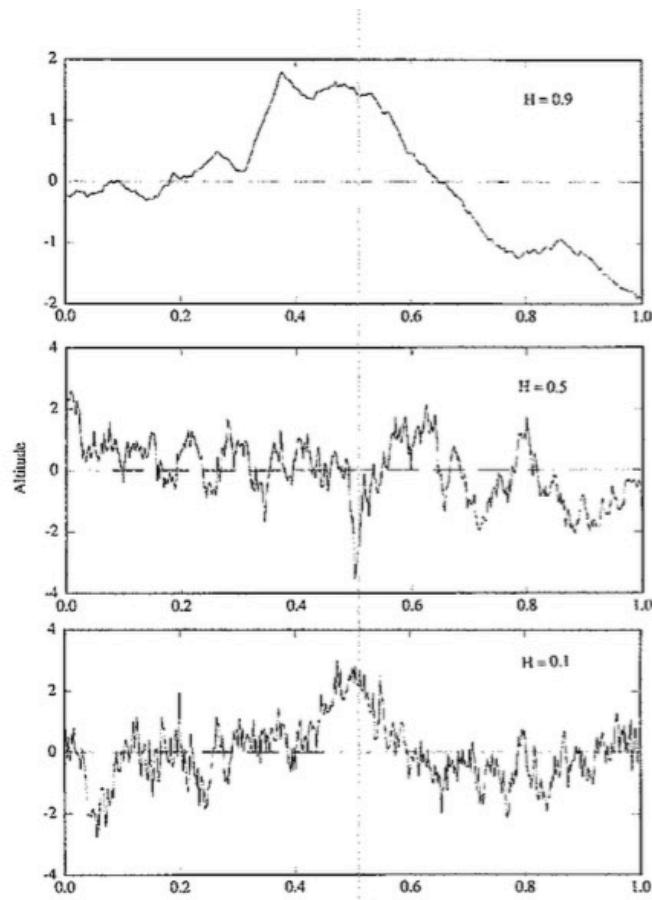


Figure 8.3: Time series with different Hurst exponents from J Feder *Fractals*, Plenum Press, 1988. Page 182

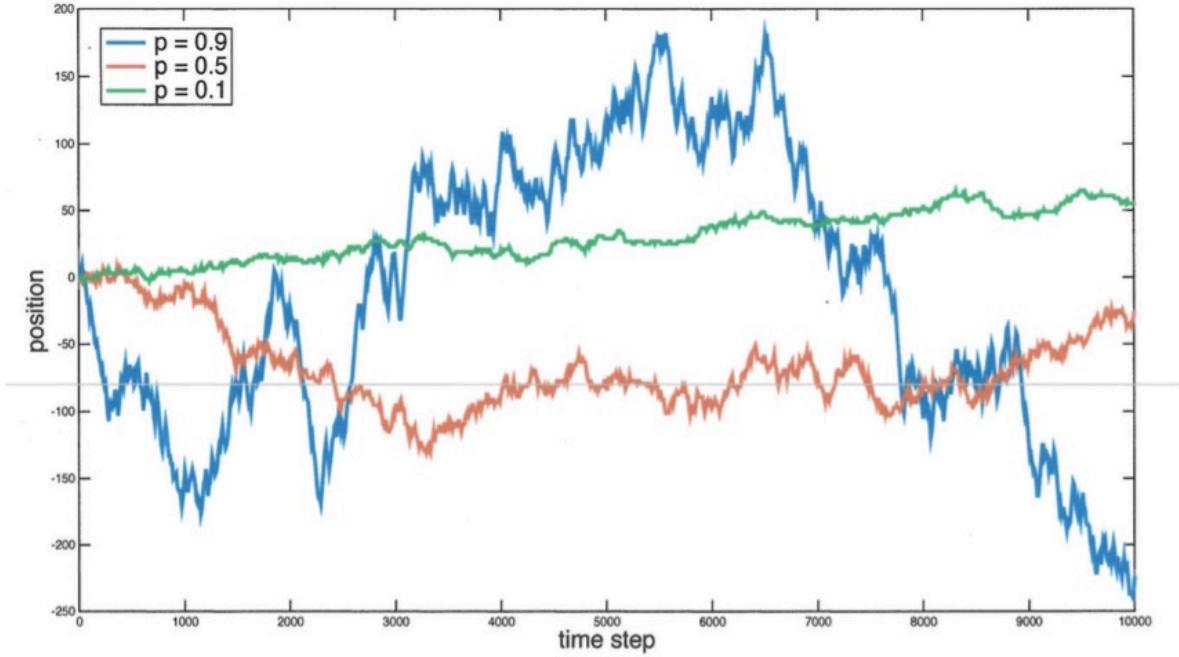


Figure 8.4: Simulated time series for the random walker defined by $X(t+1) = X(t) + \Delta(t)$ and Eq. (8.39).

8.3.1 Random walk with persistence or anti-persistence

Let

$$X(t+1) = X(t) + \Delta(t) \text{ with } \Delta(t) \in \{-1, 1\}. \quad (8.38)$$

We make the walk either persistent or anti-persistent in the following way

$$\Delta(t) = \begin{cases} \Delta(t-1) & \text{with probability } p \\ -\Delta(t-1) & \text{with probability } 1-p \end{cases} \quad (8.39)$$

For $p > 1/2$ the increment will tend to be in the same direction as the previous increment, i.e. persistence. For $p < 1/2$ the tendency will be to make an increment in the direction opposite to direction of the previous increment, so anti-persistence. And for $p = 1/2$ we have the ordinary random walk with uncorrelated increments. See Fig. 8.4 for simulated realisations of the process.

The position $X(t)$ and the increments $\Delta(t)$ become correlated because the temporal evolution of $\Delta(t)$ determines $X(t)$. The process is accordingly described by the

joint probability density $P(X, \Delta, t)$. It is straight forward to write down the Master Equation controlling the time evolution of $P(X, \Delta, t)$.

Recall that the Master Equation for the probability density for a stochastic variable $\chi(t)$ is of the form

$$P(\chi, t + 1) = P(\chi, t) + \sum_{\chi'} W(\chi, \chi') P(\chi', t) - \sum_{\chi'} W(\chi', \chi) P(\chi, t). \quad (8.40)$$

Here the probability for transitions from χ to χ' is given by $W(\chi, \chi')$ and from χ' to χ by $W(\chi', \chi)$.

Note on continuous time

For a continuous time process, one will start with

$$P(\chi, t + dt) = P(\chi, t) + [\sum_{\chi'} W(\chi, \chi') P(\chi', t) - \sum_{\chi'} W(\chi', \chi) P(\chi, t)]dt. \quad (8.41)$$

leading to

$$\frac{dP(\chi, t)}{dt} = \sum_{\chi'} W(\chi, \chi') P(\chi', t) - \sum_{\chi'} W(\chi', \chi) P(\chi, t). \quad (8.42)$$

Where now the coefficients $W(\chi, \chi')$ now denote transition probabilities per time, or rates.

To do this we need to identify the transitions, which the pair X, Δ can undergo. These are

$$\begin{aligned} (X - 1, 1) &\mapsto (X, -1) \text{ with probability } 1 - p \\ (X - 1, 1) &\mapsto (X, 1) \text{ with probability } p \\ (X + 1, -1) &\mapsto (X, -1) \text{ with probability } p \\ (X + 1, -1) &\mapsto (X, 1) \text{ with probability } 1 - p. \end{aligned} \quad (8.43)$$

Hence we arrive at the following set of coupled equations.

$$\begin{aligned} P(x, -1, t + 1) &= (1 - p)P(X - 1, 1, t) + pP(x + 1, -1, t) \\ P(x, 1, t + 1) &= pP(X - 1, 1, t) + (1 - p)P(x + 1, -1, t). \end{aligned} \quad (8.44)$$

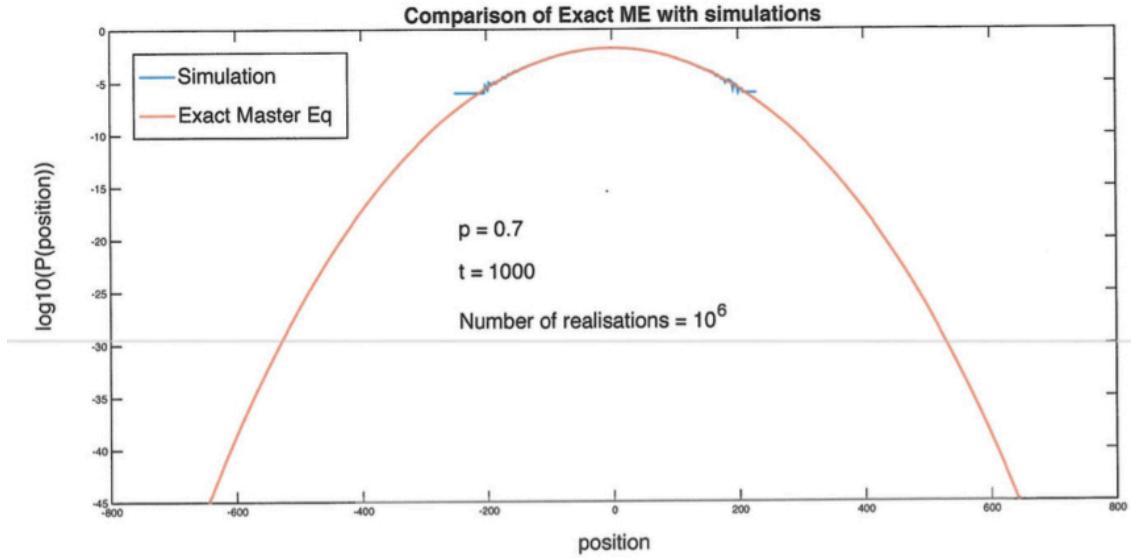


Figure 8.5: Comparison between simulation of the random walk and iterated solution of Master Equation

We have made use of the fact that the random walker always make a move with probability 1. This corresponds to cancelation between the term $P(\chi, t)$ and the term

$$\sum_{\chi'} W(\chi', \chi) P(\chi, t)$$

in Eq. (8.40).

This set of equations is easy to solve by exact numerical iteration. In Fig. 8.5 the solution of the exact Master Equation for the initial conditions $P(0, -1, 0) = p$ and $P(0, 1, 0) = 1 - p$ and $P(X, \Delta = 0) = 0$ is compared with the results of a direct simulation of the process.

Note that for an ordinary random walk with time independent increments, which in each time step makes a move to the right with probability p or move to the left with probability $1 - p$, the probability that the walker is at position X at time t evolves according to

$$P(X, t + 1) = pP(X - 1, t) + (1 - p)P(X + 1, t). \quad (8.45)$$

To obtain some analytic understanding of the role of time dependence of the increments we now analyse a mean field Master Equation for the process in Eqns. (8.38) and (8.39). The mean field approximation in this case consist in decoupling the the

probability densities for $X(t)$ and $\Delta(t)$. We use the notation $P_X(x, t)$ and $P_\Delta(\delta, t)$, where we let $x \in \mathbb{Z}$ and $\delta \in \{-1, 1\}$ denote the values assumed by the stochastic variables X and Δ , respectively. We have

$$P_X(x, t+1) = P_X(x-1, t)P_\Delta(1, t) + P_X(x+1, t)P_\Delta(-1, t) \quad (8.46)$$

and

$$\begin{aligned} P_\Delta(1, t) &= pP_\Delta(1, t-1) + (1-p)P_\Delta(-1, t-1) \\ P_\Delta(-1, t) &= (1-p)P_\Delta(1, t-1) + pP_\Delta(-1, t-1). \end{aligned} \quad (8.47)$$

We introduce the following simple notation

$$\begin{aligned} f(x, t) &:= P_X(x, t) \\ g(\Delta, t) &:= P_\Delta(\delta, t), \end{aligned} \quad (8.48)$$

to obtain the simple expressions

$$\begin{aligned} f(x, t+1) &= [pf(x-1, t) + (1-p)f(x+1, t)]g(1, t-1) \\ &\quad + [(1-p)f(x-1, t) + pf(x+1, t)]g(-1, t-1). \end{aligned} \quad (8.49)$$

By defining the matrix

$$\mathbf{M} = \begin{Bmatrix} p & 1-p \\ 1-p & p \end{Bmatrix}, \quad (8.50)$$

we can write the equation for $g(\Delta, t)$ in the form

$$\begin{pmatrix} g(1, t-1) \\ g(-1, t-1) \end{pmatrix} = \mathbf{M} \begin{pmatrix} g(1, t-2) \\ g(-1, t-2) \end{pmatrix} = \mathbf{M}^{t-1} \begin{pmatrix} g(1, 0) \\ g(-1, 0) \end{pmatrix}. \quad (8.51)$$

Here is an outline of how to solve Eq. (8.49) and Eq. (8.51).

- (1) Observe that $|g(1, t) - g(-1, t)| \rightarrow 0$ when $t \rightarrow \infty$.
- (2) Diagonalise Eq. (8.51) for $g(\Delta, t)$.
- (3) Solution

$$\begin{aligned} g(1, t-1) &= \frac{1}{2} + \beta \lambda^{t-1} \\ g(-1, t-1) &= \frac{1}{2} - \beta \lambda^{t-1}, \end{aligned} \quad (8.52)$$

where $\beta = \frac{1}{2}(g(1, 0) - g(-1, 0))$ and $\lambda = 2p - 1$.

(4) Now assume $p \leq 1/2$ and solve for $f(x, t + 1)$ in the continuum approximation

$$\partial_t f = \frac{1}{2} \partial_x^2 f - 2\beta \lambda^t \partial_x f \quad (8.53)$$

by Fourier transform in x and conclude

$$f(x, t) = \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{1}{2t}(x + h(t))^2\right], \quad (8.54)$$

where $h(t) = \gamma\tau(\exp[-t/\tau] - 1)$ with $\tau = -1/\ln(\lambda)$ and $\gamma = 2\beta$.

(5) Now derive the asymptotic behaviour for $t \rightarrow \infty$

$$\begin{aligned} \langle x \rangle(t) &\simeq -h(t) \rightarrow \gamma\tau \\ \langle x^2 \rangle(t) &= t + h^2(t) \rightarrow (\gamma\tau)^2. \end{aligned} \quad (8.55)$$

And for $t \ll$:

$$\begin{aligned} \langle x \rangle(t) &\simeq \gamma t \\ \langle x^2 \rangle(t) &\simeq t + (\gamma t)^2. \end{aligned} \quad (8.56)$$

The details of the above analysis are included in the hand write notes at the end of this chapter.

The anti-persistent case $p < 1/2$ is more difficult because the above assumed continuum approximation doesn't work.

Figure 8.6 contains a comparison of $\langle x \rangle(t)$ and $\langle x^2 \rangle(t)$ between the actual process as realised by a standard MonteCarlo simulation, the mean field master equation and the analytic solution of that equation.

Note that if we use the bottom panel of Fig. 8.6 to obtain an empirical estimate of the short time Hurst exponent defined in Eq. (8.37) we will $H \simeq 0.9$ for $p = 0.9$ and $H = 1/2$ for $p = 1/2$ and finally that $H \simeq -0.1$ for $p = 0.1$. However if we consider the long time behaviour we'll conclude that $H = 1/2$ in all three cases. This corresponds to the fact that the correlations put into the time signals through the correlations of the increments disappear with time. We would have to put in long time correlations in order to obtain a hurst exponent different from $1/2$ at long times. See e.g. [20].

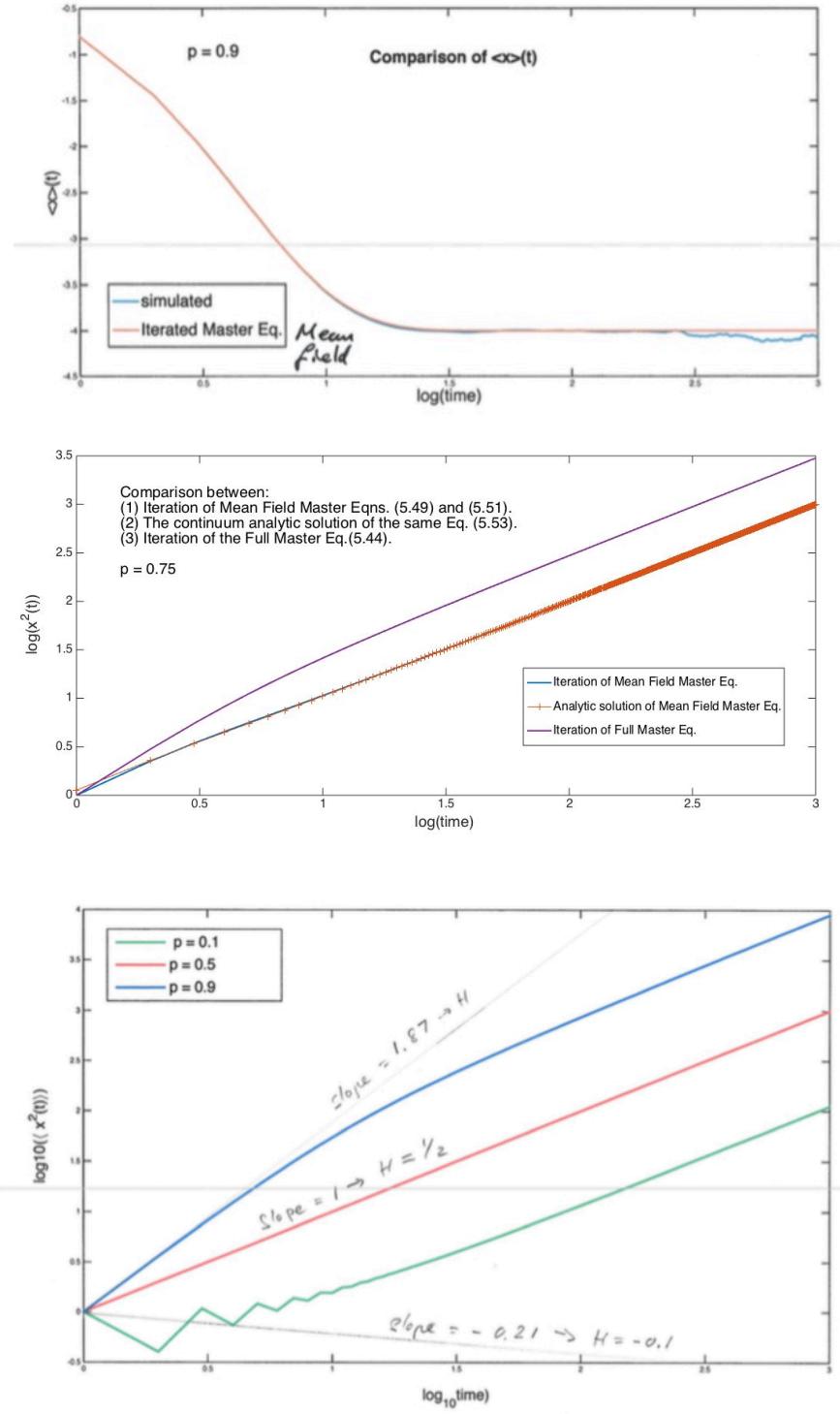


Figure 8.6: Comparison between the simulated process, the iterated mean field Master Equation, the approximative analytic solution of the latter and the iterated solution of the full Master Equation. Note that for $p = 0.9$ we have $\gamma = -0.8$ and $\tau \simeq 4.5$ and hence predict the asymptotic plateau in the top panel to be -3.6 . The curves in the bottom panel are from a direct simulation of the process.

Chapter 9

Synchronisation

Synchronisation occurs in many different situations. There exist a huge literature describing the importance of synchronisation as well as how it comes about.

- S. Strogatz, *Sync: The Emerging Science of Spontaneous Order*, Penguin, 2004.
- A. Pikovsky, M. Rosenblum and J Kurths, *Synchronization: A universal concept in nonlinear sciences*, Cambridge University Press, 2001.

As just one example of how important synchronisation is considered to be I want to mention the hypothesis by Varela et al. [74] that the state of the brain is defined in terms of its pattern of synchronisation.

Synchronisation is often taken in a broad sense. Two pendulums that interact and after a while swing in total sync, but one also often talk about synchronisation in a more loose sense. Like two temporal evolutions that follow each other to some degree. One spectacular example of partly or full synchronisation is the famous ability of certain fireflies to emit flashes of light in a synchronous way.

To study how synchronisation occurs between a number of oscillators we now discuss a very famous model introduced by Kuramoto in 1984 and discussed in the book by Pikovsky et al mentioned just above. We consider N rotors, each characterised by a phase variable ϕ_k for $k = 1, 2, \dots, N$. The time evolution of the rotors is given by

$$\frac{d\phi_k}{dt} = \omega_k + \frac{\epsilon}{N} \sum_{j=1}^N \sin(\phi_j - \phi_k). \quad (9.1)$$

The term $\sin(\phi_j - \phi_k)$ will clearly tend to make $\phi_j(t)$ and $\phi_k(t)$ equal since if, say, ϕ_k is lacking behind ϕ_j the term will be positive and therefore speed up ϕ_k . And similarly

when ϕ_k ahead of ϕ_j the sin-term will make ϕ_k slow down. However, if the frequencies ω_k are not all equal the adjustment of the speed of the rotors needed to make them synchronise will have to compete against the speed imposed by the first term ω_k of Eq. (9.1).

We see that when the ω_k assumes different values for different rotors, in order to achieve synchronisation, the rotors have to find a compromise between the speed favoured by their ω_k term and some collectively "agreed" speed, perhaps one would guess that the common speed would be given by $\bar{\omega}_k = (\sum_k \phi_k)/N$. But then this appears not to always be possible. Think of situations where the coupling ϵ is very weak and the distribution of the ω_k frequencies very broad, then some for some k we'll have that ω_k differs a great deal from $\bar{\omega}_k$ and since the first term in Eq. (9.1) dominates for when ϵ is small it is unlikely that ϕ_k can move with a velocity that significantly differs from ω_k .

From these considerations we expect that the extend of synchronisation, i.e. how many rotors that are moving with velocities that are equal, or nearly equal, will depend on the relation ship between the with, σ_g of the distribution $g(\omega)$ from which the intrinsic frequencies of the rotors are drawn and the strength ϵ of the interrelation between the rotors.

To analyse this behaviour in more mathematical detail we introduce

$$K e^{i\theta} = \frac{1}{N} \sum_{k=1}^N e^{i\phi_k}. \quad (9.2)$$

We will see below that the modulus K will act as an order parameter in the sense discussed in Sec. 10.3. Namely, K allows us to distinguish between no synchronisation $K = 0$ and some degree of synchronisation $K > 0$ and that these two situations are separated by a critical value K_c .

Next we introduce the order parameter into Eq. (9.1) in the following way

$$\begin{aligned} \frac{d\phi_k}{dt} &= \omega_k + \frac{\epsilon}{N} \sum_{j=1}^N \sin(\phi_j - \phi_k) \\ &= \omega_k + \frac{\epsilon}{N} \operatorname{Im} \sum_{j=1}^N e^{i(\phi_j - \phi_k)} \\ &= \omega_k + \frac{\epsilon}{N} \operatorname{Im} \left\{ \left(\sum_{j=1}^N e^{i\phi_j} \right) e^{-i\phi_k} \right\} \\ &= \omega_k + \epsilon \operatorname{Im} \left\{ K e^{i\theta} e^{-i\phi_k} \right\} \\ &= \omega_k + \epsilon K \sin(\theta - \phi_k). \end{aligned} \quad (9.3)$$

Consider the deviation ψ_k of rotor k from the speed given by the average of the intrinsic frequencies such that

$$\phi_k(t) = \bar{\omega}t + \psi_k(t). \quad (9.4)$$

Assume that $\theta(t) = \bar{\omega}t$. This assumption will have to be checked for consistency later, but for now we observe that substituting Eq. (9.4) into Eq. (9.3) we have

$$\frac{d\psi_k}{dt} = \omega_k - \bar{\omega} - \epsilon K \sin(\psi_k). \quad (9.5)$$

This equation allows us to study to which degree synchronisation can take place. Eq. (9.5) supports two types of solutions:

- (1) If $\psi_k(t) \rightarrow \text{constant}$ as $t \rightarrow \infty$ then we have entrainment or synchronisation of rotor k to θ .
- (2) If $\psi_k(t)$ monotonically increases (or decreases), ψ_k does not synchronise to θ .

It is straightforward to see, for example graphically – see Fig. 9.1, that those rotors for which $|\omega_k - \bar{\omega}| < \epsilon K$ case (1) above applies, and for those with intrinsic frequencies farther away from $\bar{\omega}$, i.e. $|\omega_k - \bar{\omega}| > \epsilon K$ case (2) applies.

We made the assumption $\theta = \bar{\omega}t$, we substitute this ansatz on the left hand side of Eq. (9.2) and substitute on the right hand side of this equation the expression for ϕ_k given in Eq. (9.4). From this we obtain

$$\begin{aligned} Ke^{i\bar{\omega}t} &= \frac{1}{N} = \sum_{k=1}^N e^{i(\bar{\omega}t + \psi_k)} \\ &\Downarrow \\ K &= \sum_{k=1}^N e^{i\psi_k}. \end{aligned} \quad (9.6)$$

The next step is to approximate the sum over k in Eq. (9.6) by an integral over the distribution of values assumed by the phases ψ_k , i.e. we write

$$K = \int_{-\infty}^{\infty} d\psi n(\psi) e^{i\psi}. \quad (9.7)$$

One arrives at this approximation by grouping the terms in the sum over k in Eq. (9.6) according to the value of ψ_k . Thus $n(\psi)d\psi$ denotes the number of rotors that have their ψ_k value in the interval $[\psi, \psi + d\psi]$.

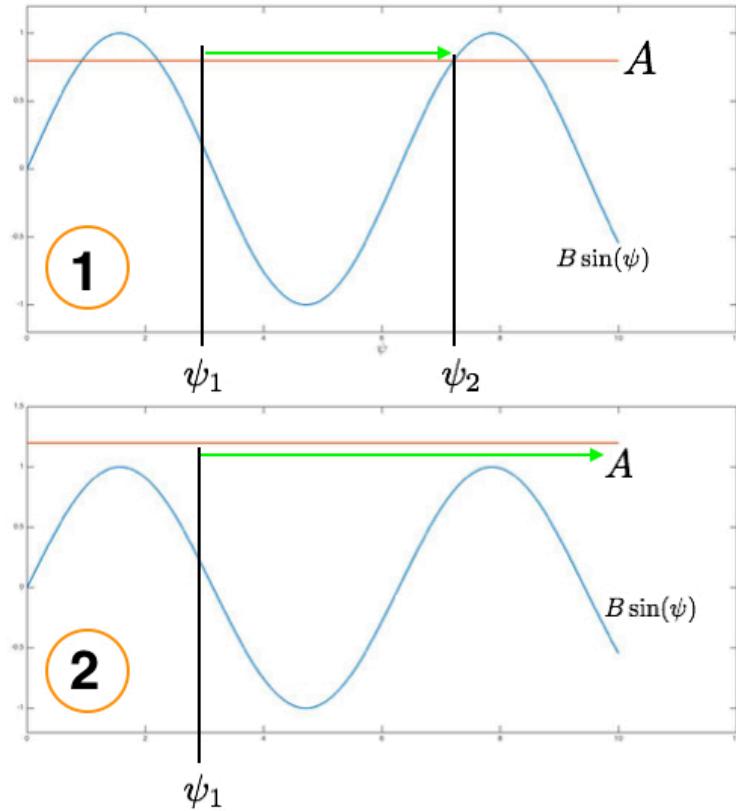


Figure 9.1: Sketch of how the relative size of $A = \omega_k - \bar{\omega}$ compare to $B = \epsilon K$ determines whether case (1) or case (2) applies. In the top panel ψ moves from ψ_1 and to ψ_2 where it stops because the right hand side of Eq. 9.5 becomes equal to zero at ψ_2 . In the bottom panel $A > B$ and therefore $\dot{\psi}$ remains positive for all ψ allowing $\psi(t)$ to monotonically increase forever.

To compute K from the integral in Eq. (9.7) we need to estimate the distribution of the phases ψ . To do this we distinguish between the rotors k that participate in synchronisation and those that don't. I.e. we need to classify the rotors according to the cases (1) and (2) above. We write

$$n(\psi) = n_s(\psi) + n_{as}(\psi)$$

where n_s denote the density of rotors that synchronise, i.e. case (1), and n_{as} denote the density of those that moves asynchronously, i.e. case (2).

For those that synchronises we have $\dot{\psi} = 0$ so according to Eq. (9.5) these rotors satisfy

$$\omega = \bar{\omega} + \epsilon K \sin(\psi).$$

This equation allows us to express ψ in terms of the ω for that rotor, the average $\bar{\omega}$ and ϵK . So we can find the density of ψ in the following way

$$n_s(\psi) = g(\omega) \left| \frac{d\omega}{d\psi} \right| = g(\bar{\omega} + \epsilon K \sin \psi) \epsilon K \cos \psi. \quad (9.8)$$

The rotors that belong to case (2), i.e. the asynchronous rotors, the phase ψ increases steadily with time and given the periodicity of Eq. (9.5) it is straight forward¹ to show that $n_{as}(\psi + \pi) = n_{as}(\psi)$. This periodicity ensures that the asynchronous rotors don't contribute to K in Eq. (9.7) since

$$n_{as}(\psi + \pi) e^{i\psi + \pi} = -n_{as}(\psi + \pi) e^{i\psi} = -n_{as}(\psi) e^{i\psi}.$$

We can therefore concentrate on the rotors that belong to case (1):

$$\begin{aligned} K &= \int d\psi n_s(\psi) e^{i\psi} \\ &= \int_{-\pi/2}^{\pi/2} d\psi g(\bar{\omega} + \epsilon K \sin \psi) \epsilon K \cos \psi e^{i\psi}. \end{aligned} \quad (9.9)$$

¹ Note that in this case $\psi(t)$ is a solution to the equation $\dot{\psi} = A - B \sin(\psi)$ with $A > B$. Integrate the equation to get

$$\begin{aligned} t &= \int \frac{d\psi}{A - B \sin \psi} \\ &= \frac{2}{\sqrt{A^2 - B^2}} \arctan \frac{A \tan \frac{\psi}{2} - B}{\sqrt{A^2 - B^2}}. \end{aligned}$$

We made use of the π periodicity to restrict the support of $n_s(\psi)$ to the interval $[-\pi/2, \pi/2]$. To proceed we need to make assumptions about how the intrinsic frequencies of the individual rotors are distributed. Well assume that the ω_k are drawn from a peaked distribution $g(\omega)$, which is symmetric about its mean:

$$g(\bar{\omega} + \omega) = g(\bar{\omega} - \omega), \quad \omega \in \mathbb{R}.$$

In this case there will be no contribution from the imaginary part of the integral in Eq. (9.9)². We are therefore lead to the following transcendental equation that determines K in terms of $\bar{\omega}$ and ϵ .

$$1 = \epsilon \int_{-\pi/2}^{\pi/2} d\psi g(\bar{\omega} + \epsilon K \sin \psi) \cos^2 \psi. \quad (9.10)$$

We'll first derive a general expression for K by expanding $g(\omega)$ to 2nd order about $\bar{\omega}$ after that we'll mention the result obtained when we assume a specific functional form of $g(\omega)$.

First expand

$$g(\bar{\omega} + \epsilon K \sin \psi) \simeq g(\bar{\omega}) + \frac{g''(\bar{\omega})}{2} (\epsilon K \sin \psi)^2. \quad (9.11)$$

Inspired by the assumed symmetry we assume $g'(\bar{\omega}) = 0$. Next, since

$$\int_{-\pi/2}^{\pi/2} d\psi \cos^2 \psi = \frac{\pi}{2} \quad \text{and} \quad \int_{-\pi/2}^{\pi/2} d\psi \sin^2 \psi \cos^2 \psi = \frac{\pi}{8},$$

²This assumption at the same time makes θ equal to the average of the time derivative of ϕ_k . Namely, we assumed $\theta(t) = \bar{\omega}t$ and from Eq. (9.4) we have $\dot{\phi}_k = \bar{\omega}t + \psi_k$. So for $\theta t = \langle \phi_k \rangle$ we must have

$$\frac{d\langle \phi_k \rangle}{dt} = 0.$$

The symmetry assumption on $g(\omega)$ ensures this, since by use of Eq. (9.5) we have

$$\begin{aligned} \langle \dot{\phi}_k \rangle &= \frac{1}{N} \sum_{k=1}^N \dot{\phi}_k \\ &= \frac{1}{N} \sum_k (\omega_k - \bar{\omega} - \epsilon K \sin \psi_k) \\ &= -\epsilon K \frac{1}{N} \sum_k \sin \psi_k = -\epsilon K \int_{-\infty}^{\infty} n(\psi) \sin \psi. \end{aligned}$$

we get

$$\epsilon g(\bar{\omega}) \frac{\pi}{2} - \frac{\epsilon^3 |g''(\bar{\omega})|}{2} K^2 \frac{\pi}{8} = 1, \quad (9.12)$$

where we explicated the assumption $g''(\bar{\omega}) < 0$. Solving this equation we get

$$\epsilon_c = \frac{2}{\pi g(\bar{\omega})} \quad (9.13)$$

and further more

$$K = \begin{cases} 0 & \text{for } \epsilon < \epsilon_c \\ \sqrt{\frac{8g(\bar{\omega})}{|g''(\bar{\omega})|\epsilon^3}} (\epsilon - \epsilon_c)^{\frac{1}{2}} & \text{for } \epsilon > \epsilon_c \end{cases} \quad (9.14)$$

As a concrete example assume

$$g(\omega) = \frac{\gamma}{\pi[(\omega - \bar{\omega})^2 + \gamma^2]}.$$

In this case $\epsilon_c = 2\gamma$ and $K = \sqrt{1 - 2\gamma/\epsilon}$ for $\epsilon > \epsilon_c$.

Chapter 10

Statistical Mechanics: From individual component to systems level

Since ancient time human enquiry has tried to make sense out of the surrounding world by relating the observed behaviour of many component system to the behaviour of the individual parts and the interaction amongst these.

Matter in the universe is organised in a hierarchical structure. At the bottom (if there is one) we have elementary particle, atoms and molecules from which we get macro molecules like proteins and DNA, these are the building blocks of organelles, which together form the cells. From cells we get organs, which put together form organisms: animals and plants of a great variety of species. One level of structure emerges from the level below. Is it possible to scientifically describe, let alone, predict emergence. Sometimes emergence is described as a phenomena beyond analysis. The perplexity with which this concept is sometimes met, is well illustrated by the following quote from a recent call for participation in a meeting held by the British research council EPSRC to look at ways to explore emergence in complex systems. Emergence is described in the following words: “For the first time since the enlightenment in the western tradition we have started to understand that there are non-causal systems in which some things just *are*. The concept of emergence by which patterns of possibility arise through interactions of agents over time, accepts that even with the same starting conditions the same pattern would not necessarily repeat.” Another attitude is represented by Lord Roberts May’s statement in a resent lecture that *when* people say something is an emergent property, it just means they don’t understand the phenomena.

We will argue that emergence is neither an empty concept nor a mysterious non-causal enigma. On the contrary, emergence is central to scientific enquiry. Emergence occurs when many components interact and combine to form an identifiable system. In philosophy this is the observation that quantitative changes accumulate and give rise to new qualitative changes. A proposition that can be traced from the ancient Greek philosophers through Hegel to Dialectical Materialism. In physics Phil Anderson famously summed the fact that new levels of organisation need new types of description, up in the phrase *More is Different*[3]. By following the tradition of statistical mechanics it is sometimes possible to reduce significantly the confusion surrounding what emergence is and how it can be investigated and described. Kenneth Wilson got the 1982 Nobel prize for his Renormalisation Group theory, which is a particular beautiful method for extracting certain emergent properties with great mathematical detail and precision[78].

Statistical mechanics is concerned with the interaction of many components. From interactions between the components at one given level the aim is to understand the collective coherent behaviour, which emerges as many components are put together. It is through the interactions of the components at one level that the next level emerges. We consider a collection of *interacting* atoms and the outcome is, say, a transistor. In some cases the microscopic details of the properties of the *individual* building blocks are not so crucial. It happens that the collective behaviour is controlled by general properties of the *interactions* between the components more than by the intrinsic properties of the components. A number of methods have been developed to bridge the gap between the individual components comprising a system and the collective whole. This often involves predictions of the asymptotic behaviour at long distances and at long times. In particular the philosophy and technique of the Renormalisation Group have been successful in a number of cases in doing this. Many other approaches exist, sometimes a coarse grained description is sought similar to the one used in fluid dynamics. This relates to mean field descriptions in which one focus on how the average behaviour.

Since essentially any activity within statistical mechanics is concerned with a description of emerging phenomena a very large literature exists and we list here only a few books of particular relevance to the view point of the present article [59, 44, 4, 38, 8, 49, 12, 73, 70, 10, 62].

10.1 How to obtain the probability measures

To describe how statistical mechanics is able to identify structures emerging at the macroscopic level we briefly recall how macroscopic (or systems level) quantities are

obtained through averaging procedures. The reason equilibrium systems can be analysed in particular detail is that the situation, where the systems of interest can be considered as in thermal equilibrium with a heat bath, allows for the determination of the probability weights of the individual micro-states. One starts out with the following fundamental hypothesis concerning isolated or closed systems:

- **Micro Canonical Ensemble**

For a closed system it is assumed that *all* micro-states, consistent with the macroscopic constraints, occur with *equal* probability.

The macroscopic constraints can, for example, be the total energy E (which is constant for a closed system) and the volume V . Denote by $\Omega(E, V)$ the total number of micro-states possible under these constraints. Meaning the components or particles of the system have to be located within the given volume V , when we add all the energies of the particles the sum must equal E . The probability $p(s)$ that the system is in a particular state s is then

$$p(s) = \frac{1}{\Omega(E, V)}. \quad (\text{10.1})$$

Closed systems are not very interesting in the sense that one is unable to interact with them. A much more interesting situation is when the system **S** under consideration is brought in contact with a heat bath **B** or heat reservoir. The heat bath is a system so big that even when it exchanges energy with the small system of experimental interest, the heat bath remains unchanged. Say a cup of tea in contact with the Pacific Ocean. The heat bath is characterised by its temperature T . We can now use the fundamental hypothesis above to determine the probabilistic weights for the states of **S**. Since the combined system **B + S** is closed the weights for the combined system is given by the Micro Canonical Ensemble, i.e. all micro-states of the combined system are equally likely. The number of micro-states for the combined system of total energy $E_{Tot} = E_B + E_S$ will be a product

$$\Omega_{Tot}(E_{Tot}) = \Omega_B(E_B)\Omega_S(E_S). \quad (\text{10.2})$$

Here one neglects the interactions between the heat bath and the system. Now focus on one particular micro-state s of **S** of energy E_s . Since we have a particular state, s , in mind we have $\Omega_S(s) = 1$. This state can be combined in many ways with states of the bath **B** as long as those fulfil the constraint $E_{Tot} = E_B + E_s$. So the probability, $p(s)$, for finding the system **S** in s , when **S** is in equilibrium with the bath, is proportional to $\Omega_B(E_{Tot} - E_s)$. Namely

$$p(s) = \frac{\Omega_B(E_{Tot} - E_s)}{\sum_{state} \Omega_B(E_{Tot} - E_{state})}, \quad (\text{10.3})$$

the denominator ensures normalisation. In order to introduce the temperature into the mathematical formalism it turns out that we should consider the logarithm of $p(s)$. We have

$$\ln[p(s)] = \text{constant} + \ln[\Omega_{\mathbf{B}}(E_{Tot} - E_s)] \quad (10.4)$$

$$= \text{constant} + \ln[\Omega_{\mathbf{B}}(E_{Tot})] - \frac{\partial \ln[\Omega_{\mathbf{B}}(E_{Tot})]}{\partial E_{Tot}} E_s \quad (10.5)$$

$$= \text{constant} - \frac{1}{k_B T} E_s. \quad (10.6)$$

Here we Taylor expanded to linear order to obtain the first equality. The second equality follows, because it can be shown by use of the first and second law of thermodynamics that the temperature is given by

$$\frac{1}{k_B T} = \frac{\partial \ln[\Omega_{\mathbf{B}}(E_{Tot})]}{\partial E_{Tot}}. \quad (10.7)$$

This is obtained in the following way. The first and second law of thermodynamics lead to the following thermodynamic identity $dE = TdS - pdV$ where the entropy $S = k_B \ln[\Omega(E)]$. Since the thermodynamic identity takes the form of an exact differential we conclude that $\partial E / \partial S = T$ from which Eq. (10.7) follows. We now conclude

$$p(s) = \frac{e^{-\frac{E_s}{k_B T}}}{Z}, \quad (10.8)$$

where the constant Z , called the partition function, is obtained from the normalisation condition

$$\sum_{\text{states}} p(s) = 1, \quad (10.9)$$

to be given by

$$Z = \sum_{\text{states}} e^{-\frac{E_s}{k_B T}} = \sum_{\text{states}} e^{-\beta E_s}, \quad (10.10)$$

where we introduced the standard notation $\beta = 1/k_B T$.

The partition function plays a central role in statistical mechanics because statistical average quantities can be obtained from it. For example the average energy $\langle E \rangle$ is given by

$$\begin{aligned}
\langle E \rangle &= \sum_{\text{states}} E_S p(s) \\
&= \sum_{\text{states}} E_S \frac{e^{-\frac{E_S}{k_b T}}}{Z} \\
&= -\frac{\partial \ln(Z)}{\partial \beta}.
\end{aligned} \tag{10.11}$$

We conclude that the probabilistic weights, needed to calculate the average macroscopic behaviour of a systems in contact with a heat bath at temperature T , is given by the (Boltzmann) weights in Eq. (10.8). And we mention that a large number of average quantities can be calculated from the sum in Eq. (10.10). This important sum is called the *partition function* or partition sum. Some states, or configurations of the microscopic degrees of freedom, will contribute more to the partition sum than others, such configurations can sometimes be identified as macroscopic collective excitations. These may possess a degree of robustness and stability and can in such cases be identified as macroscopic emergent objects with specific properties that can be considered essential building blocks. Perhaps it is instructive to have the following picture in mind. Think of a pool table. To describe the motion of the balls we can either follow the trajectories of all the individual molecules making up 15 coloured balls or we can notice that some of the molecules move together in a coordinated way and thereby form each of the 15 balls. We can therefore instead simply follow the trajectories of the centre of mass (COM) of each of the balls. Obviously we lose a lot of information since we can't go from the COM of the balls to the motion of all the molecules; whereas we can drive the COM motion if we know the motion of all the molecules. Hence we note that emergence involves a loss of information. We will discuss in detail an important and illustrative example in the next section.

10.2 Ising model and mean field

The Ising model originates in physics where it is a model of magnetism. The magnetism of a material is produced by the alignment of atomic magnetic moments and is fundamentally a quantum phenomena. Nevertheless, it has turned out that many macroscopic properties of magnetic materials can be understood by use of very simply statistical mechanics models that do not explicitly refer to quantum mechanics. We'll introduce the Ising model and describe how it is easily handled in the mean field approximation. The model considers N degrees of freedom S_i with $i = 1, \dots, N$. These

are called spin and we assume they are placed on a d dimensional hypercube. See Fig. 10.1 for a two dimensional sketch.

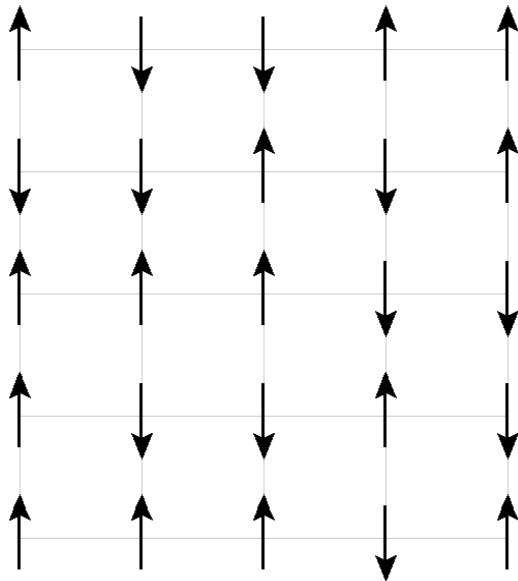


Figure 10.1: Sketch of two dimensional Ising model

Each spin can assume two values, say up $S_i = 1$ and down $S_i = -1$. The model is defined by the energy assigned to each configuration of the spin variables which is given by Ising Hamiltonian

$$H[S_1, \dots, S_N] = -\frac{1}{2}J \sum_{i=1}^N \sum_{j \in \text{nb}(i)} S_i S_j - B \sum_{i=1}^N S_i \quad (10.12)$$

Here J describes the coupling between the spins, the set of neighbours of spin i is denoted by $\text{nb}(i)$ and B the coupling of the individual spins to an external applied

magnetic field. Our aim is to compute the macroscopic magnetization which is given by the average direction of an arbitrary spin S_k :

$$\langle S_k \rangle = \sum_{S_1 \in \{-1,1\}} \dots \sum_{S_N \in \{-1,1\}} S_k \frac{e^{-\beta H[S_1, \dots, S_N]}}{Z} \quad (10.13)$$

Due the first term in the Hamiltonian in Eq. (10.12), the partition function and the average in Eq. (10.13) are difficult to compute for arbitrary dimension. For $d = 1$ the calculation is straight forward, for $d = 2$ an exact result can be obtained through a fairly complicated computation, first done by Lars Onsager in 1944. No exact solution exist in higher dimensions. But we may obtain a qualitative understanding of the behaviour of the magnetization by use of a mean filed approach.

To do so we rewrite the Hamiltonian in the following way:

$$\begin{aligned} H[S_1, \dots, S_N] &= -\frac{j}{2} \sum_{i=1}^N \sum_{j \in \text{nb}(i)} S_i S_j - B \sum_{i=1}^N S_i \\ &= \sum_{i=1}^N -(\frac{J}{2} \sum_{j \in \text{nb}(i)} S_j + B) S_i \\ &:= -\sum_{i=1}^N \tilde{B}_i S_i. \end{aligned} \quad (10.14)$$

We can now obtain significant simplification by neglecting the dependence of \tilde{B}_i on i . The rationale behind this goes as follows. Let $q = |\text{nb}(i)|$ denote the number of neighbours of each spin. We now write

$$\sum_{j \in \text{nb}(i)} S_j = q \frac{\sum_{j \in \text{nb}(i)} S_j}{q} = q \bar{S}_i, \quad (10.15)$$

where \bar{S}_i denotes the average value of the spins in the neighbourhood of spin number i . If we make the assumption that this average doesn't vary much from spin to spin then we can neglect the dependence of \tilde{B}_i on i and from Eq. 10.14 we get the mean field Hamiltonian

$$H_{MF}[S_1, \dots, S_N] = -\sum_{i=1}^N \tilde{B} S_i. \quad (10.16)$$

In this approximation Eq. (10.13) becomes

$$\langle S_k \rangle = \frac{1}{Z} \prod_{i \neq k} \left(\sum_{S_i=\pm 1} e^{-\beta \tilde{B} S_i} \right) \left(\sum_{S_k=\pm 1} S_k e^{-\beta \tilde{B} S_k} \right). \quad (10.17)$$

The partition function is in this mean field approximation given by

$$Z = \prod_{i=1}^N \left(\sum_{S_i=\pm 1} e^{-\beta \tilde{B} S_i} \right) \quad (10.18)$$

and we arrive at the following simple expression

$$\langle S_k \rangle = \frac{e^{\beta \tilde{B}} - e^{-\beta \tilde{B}}}{e^{\beta \tilde{B}} + e^{-\beta \tilde{B}}} = \tanh(\beta \tilde{B}). \quad (10.19)$$

Of course this equation isn't in a closed form that allow us to compute $\langle S_k \rangle$. But if we now make the assumption that the local average \bar{S}_i and the thermodynamic average $\langle S_k \rangle$ can be considered equal than Eq. (10.19) becomes an implicit equation for $\langle S_k \rangle$:

$$\langle S \rangle = \tanh[\beta(\frac{Jq}{2}\langle S \rangle + B)]. \quad (10.20)$$

Now solve the equation to obtain $\langle S \rangle$. First consider $J = 0$, i.e. no coupling between the degrees of freedom, in this case $\langle S \rangle = \tanh[\beta B]$ and the magnetic moment is directly induced by the applied field. For $J > 0$ and $B > 0$ it is easy to see graphically that only one solution $\langle S \rangle > 0$ exists. Finally for $J > 0$ and $B = 0$ one finds easily graphically that a non-zero solution appears when the temperature becomes lower than $T_c = \frac{Jq}{2k_B}$. The fact that a preferred direction occurs at low temperature is called symmetry breaking. The symmetry between up $S_i = 1$ and down $S_i = -1$ is broken. For $B = 0$ the symmetry breaking is a collective phenomena relating to the interaction between neighbour spins. Such collective phenomena that are able to change the systemic state are often encountered in complex systems. A prominent example from sociology is opinion formation. The temperature T_c where $\langle S \rangle$ becomes non-zero is denoted the critical temperature and the system is said to be critical - or at a critical point - when $T = T_c$. The behaviour at or near critical has been a great inspiration to the theory of complex systems. We'll return to how and why this is multiple times during the course.

10.3 The peculiar nature of a critical point

Near a critical point the dependence of observables on control parameters typically takes the form of power laws. An example of this is obtained from Eq. (10.20) if we expand the hyperbolic tangent: $\tanh(x) \simeq x - x^3/3$, which leads to the expression

$$\langle S \rangle \simeq \sqrt{3}(1 - \frac{T}{T_c})^{1/2} = \sqrt{3}(1 - \frac{T}{T_c})^\beta. \quad (10.21)$$

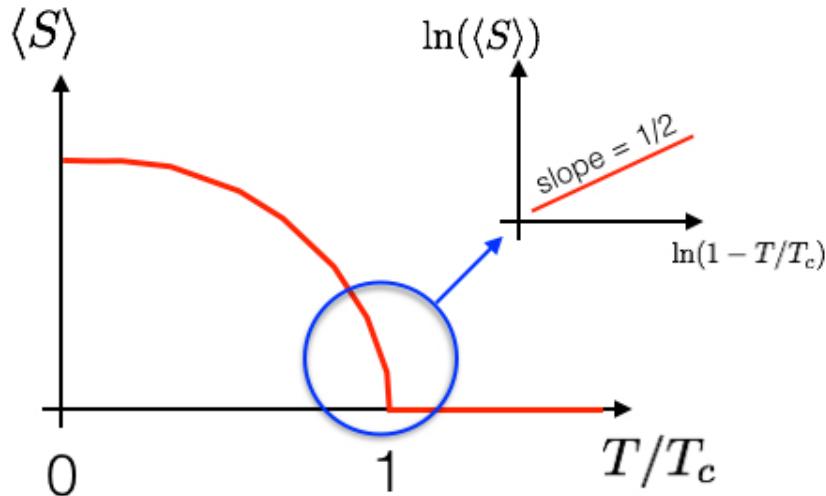


Figure 10.2: Sketch of the behaviour of the order parameter near a critical point

| d | MF $T_c [J/k_B]$ | Actual $T_c [J/k_B]$ | MF β | Actual β |
|-----|------------------|----------------------|------------|----------------|
| 1 | 1 | 0 | 1/2 | - |
| 2 | 2 | 2.269 | 1/2 | 1/8 |
| 3 | 3 | 4.52 | 1/2 | 0.326 |
| 4 | 4 | 6.68 | 1/2 | 1/2 |

Table 10.1: Critical properties of the hypercubical Ising model

This behaviour is sketched in Fig. 10.2. The exponent $\beta = 1/2$ is an example of a critical exponent. These are often difficult to compute and mean field only gives the exact result when the dimension of the system is sufficiently high. See Table 10.1. An intuitive way to understand why the mean field estimate might become correct in high dimensions is to note that the number of neighbours increase with dimension and accordingly the replacement of \bar{S}_i by \bar{S} and then by $\langle S \rangle$ is likely to be more accurate in higher dimensions. The difference between the mean field estimate and the exact behaviour is due to fluctuations away from the average behaviour. These fluctuations are stronger in lower dimensions and can be so strong as to prevent the ordering, $\langle S \rangle > 0$ to be able to occur at any $T > 0$. This is what happens for the Ising model in one dimension.

In general one defines a lower (d_L) and a higher (d_H) critical dimension. The lower critical dimension is the dimension at and below which no phase transition occurs. And the higher critical dimension is the dimension where mean field theory starts to predict the correct values of the critical exponents. For the Ising model we have

$d_L = 1$ and $d_H = 4$.

The critical point at T_c is singular in several ways. As one approach T_c one often encounter power law dependence, response functions describing how the system respond to external perturbations diverge at T_c . Right at T_c the system lack characteristic scales, which show up as power law dependence of spatial and temporal correlation functions.

The response to external perturbations is e.g. described by the heat capacity

$$C = \frac{\partial \langle E \rangle}{\partial T}, \quad (10.22)$$

and the magnetic susceptibility

$$\chi = \frac{\partial \langle M \rangle}{\partial B} = N \frac{\partial \langle S \rangle}{\partial B}. \quad (10.23)$$

Here $\langle E \rangle$ is the thermodynamic average of the total energy and $\langle M \rangle$ the thermodynamic average of the total magnetisation $M = \sum_i S_i$. The heat capacity C relates to how the energy changes if the temperature is changed a small amount and the magnetic susceptibility relates to how much the magnetization changes if the applied magnetic field is changed by a small amount.

We have in the canonical ensemble Eq. (10.11) for $\langle E \rangle$ so

$$\langle E \rangle = -\frac{\partial \ln(Z)}{\partial \beta}. \quad (10.24)$$

and according to the Hamiltonian in Eq. (10.12) we have

$$\langle M \rangle = \frac{1}{\beta} \frac{\partial \ln(Z)}{\partial B}. \quad (10.25)$$

From these expressions it is easy to derive the following very similar expressions for the two response functions.

$$C = k_B \beta^2 (\langle E^2 \rangle - \langle E \rangle^2) \quad (10.26)$$

and

$$\chi = \beta (\langle M^2 \rangle - \langle M \rangle^2). \quad (10.27)$$

we notice that in both cases the response is proportional to the variance of the relevant observable. As T_c is approached both C and χ exhibit singular behaviour, often they diverge according a power law

$$C \propto |T - T_c|^{-\alpha} \text{ and } \chi \propto |T - T_c|^{-\gamma}. \quad (\text{10.28})$$

We can easily compute the divergence of the susceptibility in the mean field approximation by differentiation of Eq. (10.20). Let $x = \langle M \rangle = N \langle S \rangle$. Here N denotes the number of sites. By implicit differentiation of Eq. (10.20) we get

$$\chi = \frac{1}{N} \frac{\partial x}{\partial B} = \left(\beta \frac{Jq}{2N} \frac{\partial x}{\partial B} + B \right) \cosh^{-2} \left(\beta \left(\frac{Jq}{2N} + B \right) \right). \quad (\text{10.29})$$

We are interested in the limit $B \rightarrow 0$ and after moving factors around obtain

$$\chi = \frac{N\beta}{\cosh^2 \left(\frac{Jq\beta}{2} \langle S \rangle \right) - \frac{Jq\beta}{2}}. \quad (\text{10.30})$$

Recall that $T_c = \frac{Jq}{2k_B}$, we will expand about $T = T_c$ and make use of Eq. (10.21) together with $\cosh(x) = 1 + x^2/2! + \dots$ to derive that

$$\chi \simeq \frac{N}{2k_B} (T_c - T)^{-1} \text{ for } T \rightarrow T_c. \quad (\text{10.31})$$

From this we observe that one way in which the critical point is special is that the fluctuations grow without bound as one approaches criticality. The singular behaviour at T_c is also reflected in the behaviour of the so called two point correlation function

$$\begin{aligned} C(t, t) &= \langle S(\mathbf{r}_0, t_0) S(\mathbf{r}_0 + \mathbf{r}, t_0 + t) \rangle - \langle S(\mathbf{r}_0, t_0) \rangle \langle S(\mathbf{r}_0 + \mathbf{r}, t_0 + t) \rangle \\ &= \langle S(\mathbf{r}_0, t_0) S(\mathbf{r}_0 + \mathbf{r}, t_0 + t) \rangle - \langle S(\mathbf{r}_0, t_0) \rangle^2 \end{aligned} \quad (\text{10.32})$$

for a system homogenous in space and time. Away from the critical point the typical leading functional dependence of $C(r, t)$ is of the form

$$C(r, t) \propto \exp(-r/\xi) \exp(-t/\tau). \quad (\text{10.33})$$

So exponential decay of correlations in space and time. At the critical point T_c the functional form changes to power law dependencies schematically given by

$$C(r, t) \propto \frac{1}{r^{2-d+\eta}} \frac{1}{t^\alpha}. \quad (\text{10.34})$$

^oActually the variance, but all of the statistical mechanics literature denotes this object as the correlation function

The change from exponential decaying correlations to power law decay comes about through the divergence of the characteristic length ξ , called the correlation function and of the characteristic time scale τ , called the correlation time, namely

$$\xi \propto |T - T_c|^{-\nu} \text{ and } \tau \propto \xi^z. \quad (10.35)$$

For the one dimensional Ising chain it is fairly straight forward to analyse this behaviour, see e.g. F. Schwabl Statistical Mechanics, Springer 2002, Appendix F. We will return to this behaviour in the context of various models of relevance to complex systems. In particular the exponential decay of spatial correlations can be related to a two value signal that switches at random between the two values. We'll return to this when we discuss time signals.

Chapter II

Emergence: From Micro to Macro - Vortices as a case study

II.1 The Two Dimensional XY-Model

Here we describe how the averaging procedure described in the previous section can be structured in a way that allows the introduction of new effective collective degrees of freedom. These describe macroscopic excitations created by the coherent motion of huge number of microscopic variables. When the collective degrees have been identified information concerning the detailed motion of the microscopic variables can be neglected and one is in this way able to reduce the computational effort needed and at the same time identify the essential emergent structures. A particular clear example of this procedure consist in the physics of systems modelled by the so called two dimensional XY-model.

We start our discussion by considering the formation of vortices in the sea of two dimensional magnetic moments. The individual microscopic magnetic moments sit on the sites of a two dimensional square lattice and the direction of the moments are confined to two dimensions see Fig. II.1 (A beautiful online interactive simulation can be found on Hans Weber's web page at <http://www.mt.luth.se/weber/>). Each magnetic moment can be thought of as a magnetic needle, or an arrow, confined to two dimensions and pointing in a specific direction described by the angle Θ . The magnetic moment number i is given by the vector $\mathbf{S}_i = (\cos(\theta_i), \sin(\theta_i))$.

We will use it as our reference model. We think of the model as consisting of planar rotors of unit length arranged on a two dimensional square lattice. The Hamiltonian

of the system is given by

$$H = -J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j). \quad (\text{II.1})$$

Here J is the coupling constant between the magnetic moments, $\langle i,j \rangle$ denotes summation over all nearest neighbour sites in the lattice, and θ_i denotes the angle of the rotor on site i with respect to some (arbitrary) polar direction in the two dimensional vector space containing the rotors.

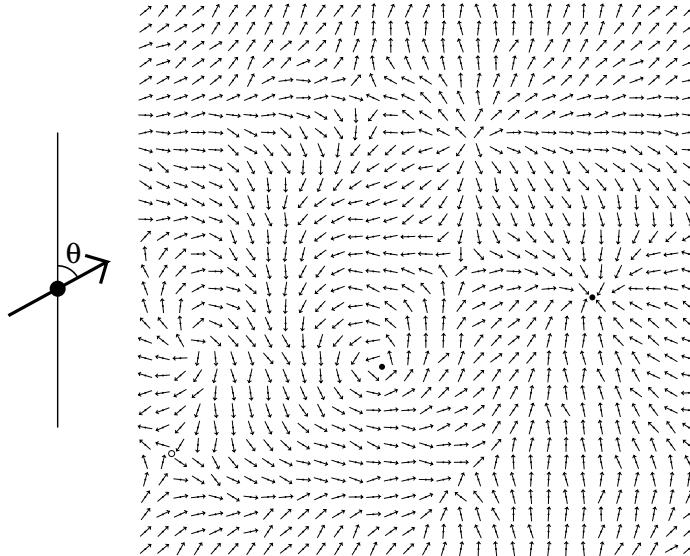


Figure II.1: Rotor configuration of the XY-model. Vortices (black circle) and anti-vortices (white circle) are clearly seen. The configuration is neutral, i.e. there is an equal number of vortices and anti-vortices; but only one circle is indicated. The reader may find it amusing to try to locate the missing anti-vortex.

We shall see below how the components, the rotors, work together to form certain collective coherent structures: topological defects or topological charges. In Fig. II.1

these excitations are depicted. Each consists of a whirl or vortex in the configuration of the rotors. There are vortices of opposite sign. As one move around in positive direction along a contour encircling the centre of a vortex (black circle) the rotors perform a full rotation in the positive direction as well. When we move around one of the anti-vortices (white circle) in a similar way, the rotors undergo a full rotation in the negative direction. Although these charges are here seen as arising from rotors or magnetic moments, the impressive fact is that these topological charges also represent Coulomb charges in two dimensions, or dislocations in two dimensional crystals, or vortices in two dimensional superconductors or a large number of other collective excitations. The interaction between the topological charges depends in all cases logarithmically on the spatial separation and this leads to some very general collective behaviour, most spectacular the logarithmic dependence on separation causes a certain type of phase transition: the Kosterlitz-Thouless transition [40].

If we assume that the direction of the rotors varies smoothly from site to site, we can approximate $\cos(\theta_i - \theta_j)$ by the first two terms $1 - \frac{1}{2}(\theta_i - \theta_j)^2$ in the Taylor expansion of cos. The sum over the nearest neighbours corresponds to the discrete Laplace operator, which we can express in terms of partial derivatives through $\theta_i - \theta_j = \partial_x \theta$ for two sites i and j which differs by one lattice spacing in the x -direction. This leads to the continuum Hamiltonian

$$H = E_0 + \frac{J}{2} \int d\mathbf{r} (\nabla \theta)^2. \quad (\text{II.2})$$

Here $E_0 = 2JN$ is the energy of the completely aligned ground state of N rotors.

The thermodynamics of the system is obtained from the partition function

$$Z = e^{-\beta E_0} \int D[\theta] \exp\left\{-\beta \frac{J}{2} \int d\mathbf{r} (\nabla \theta)^2\right\}, \quad (\text{II.3})$$

a functional integral over all possible configurations of the director field $\theta(\mathbf{r})$. Not all configurations will be of the same importance. By focusing on the terms in the sum that contribute most, we can identify the configurations that may be used as building blocks at the next level of description. Since the energy appears in the exponential with a negative sign in front, the most significant contributions will be those with the smaller energy - thus we have to pick out the local minima. We therefore divide the integral over $\theta(\mathbf{r})$ into a sum over the local minima θ_{vor} of $H[\theta]$ plus fluctuations θ_{sw} around the minima

$$Z = e^{-\beta E_0} \sum_{\theta_{vor}} \int D[\theta_{sw}] \exp\left\{-\beta(H[\theta_{vor}] + \frac{1}{2} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \theta_{sw}(\mathbf{r}_1) \frac{\delta^2 H}{\delta \theta(\mathbf{r}_1) \delta \theta(\mathbf{r}_2)} \theta_{sw}(\mathbf{r}_2))\right\}. \quad (\text{II.4})$$

The field configurations corresponding to local minima of H are solutions to the extremal condition

$$\frac{\delta H}{\delta \theta(\mathbf{r})} = 0 \Rightarrow \nabla^2 \theta(\mathbf{r}) = 0. \quad (\text{II.5})$$

There are two types of solutions to this equation. The first consists of the ground state $\theta(\mathbf{r}) = \text{constant}$. The second type of solutions consist of vortices in the director field (see Fig. 1) and are obtained by imposing the following set of boundary conditions on the circulation integral of $\theta(\mathbf{r})$:

- 1) For all closed curves encircling the position \mathbf{r}_0 of the centre of the vortex

$$\oint \nabla \theta(\mathbf{r}) \cdot d\mathbf{l} = 2\pi n, \quad n = 1, 2, \dots. \quad (\text{II.6})$$

- 2) For all paths that don't encircle the vortex position \mathbf{r}_0

$$\oint \nabla \theta(\mathbf{r}) \cdot d\mathbf{l} = 0. \quad (\text{II.7})$$

Condition 1) imposes a singularity in the director field. Note the circulation integral *must* be equal to an integer times 2π since we circle a closed path and therefore $\theta(\mathbf{r})$ has to point in the same direction after traversing the path as it did when we started.

We can estimate the energy of a vortex in the following way. The problem is spherical symmetric, hence the vortex field θ_{vor} must be of the form $\theta(\mathbf{r}) = \theta(r)$. The dependence on r can be found from Eq. (II.6). We calculate the circulation integral along a circle of radius r centred at the position \mathbf{r}_0 of the vortex

$$2\pi n = \oint \nabla \theta(\mathbf{r}) \cdot d\mathbf{l} = 2\pi r |\nabla \theta|. \quad (\text{II.8})$$

We solve and obtain $|\nabla \theta(r)| = n/r$. Substitute this result into the Hamiltonian Eq. (II.2)

$$E_{vor} - E_0 = \frac{J}{2} \int d\mathbf{r} [\nabla \theta(\mathbf{r})]^2 \quad (\text{II.9})$$

$$= \frac{Jn^2}{2} \int_0^{2\pi} d\phi \int_a^L r dr \frac{1}{r^2} \quad (\text{II.10})$$

$$= \pi n^2 J \ln\left(\frac{L}{a}\right). \quad (\text{II.11})$$

Here a denotes the lattice constant and L is the linear size of the considered lattice. The circulation condition Eq. (II.6) creates a distortion in the phase field $\theta(\mathbf{r})$ that

persists infinitely far from the centre of the vortex. $|\nabla\theta|$ decays only as $1/r$ leading to a logarithmic divergence of the energy. Hence we need to take into account that the integral over r in Eq. (II.10) is cut-off for large r -values by the finite system size L and for small r -values by the lattice spacing a . We recall that our continuum Hamiltonian is an approximation to the lattice Hamiltonian in Eq. (II.1). A vortex with the factor n in Eq. (II.6) larger than one is called multiple charged. We notice that the energy of the vortex is quadratic in the charge. In an macroscopically large system even the energy of a single charge vortex will be large, and therefore we do not expect individual vortices be thermally induced.

Consider now a pair consisting of a single charged vortex and a single charged anti-vortex. When we encircle the vortex, we pick up $\oint d\mathbf{l} \cdot \nabla\theta = 2\pi$ and when we encircle the anti-vortex, we pick up $\oint d\mathbf{l} \cdot \nabla\theta = -2\pi$. Hence, if we choose a path large enough to enclose both vortices, we pick up a circulation of the phase equal to $2\pi + (-2\pi) = 0$. I.e. the distortion of the phase field $\theta(\mathbf{r})$ from the vortex-anti-vortex pair is able to cancel out at distances from the centre of the two vortices large compared to the separation R between the vortex and the anti-vortex, see Fig. 1. This explains why the energy of the vortex pair is of the form [II, 76, 47]

$$E_{2vor}(R) = 2E_c + E_1 \ln(R/a). \quad (\text{II.12})$$

Where E_c is the energy of a vortex core and E_1 is proportional to J . In detail, the phase field $\theta_{2vor}(\mathbf{r})$ of a vortex located at $\mathbf{r} = (-a, 0)$ and an anti-vortex located at $\mathbf{r} = (a, 0)$ is given by [II]

$$\theta_{2vor}(\mathbf{r}) = \operatorname{arctg} \left(\frac{2ay}{a^2 - r^2} \right). \quad (\text{II.13})$$

Significant aspects of the macroscopic behaviour of the XY-model can be understood by treating the vortices as particles characterised by their position and their charge and ignoring the underlying sea of rotors. Indication of this follows from the expression for the energy of a pair of vortices in Eq. II.12. This energy is given in terms of the relative position of the two vortices no reference is needed to the microscopic rotor field given in Eq. II.13. The pairs of vortices has dramatic effects on the macroscopic behaviour of the XY-model. At low temperature the vortices are organised fairly small bound pairs, as the temperature is increased and more thermal energy is available the separation between paired up vortices grow and at a certain temperature the pairs break apart with the effect that the individual vortices now can move freely around as they are no longer kept in check by their partner of the opposite charge. The result is the Kosterlitz-Thouless transition which manifest itself in various ways in different

realisations of the XY-model. Before we discuss this transition we will look at the average ordering of the rotors. This quantity - the magnetisation - is usually able to monitor if a dramatic change in the macroscopic behaviour occur as function of temperature. But not so in the 2d XY-model. To understand this makes it more clear how important it is to identify correctly the emergent excitations of a many component system.

II.2 Lack of Ordering in Two Dimensions

In order to highlight the peculiarity of two dimensions we consider the d -dimensional XY-model. We imagine a d -dimensional cubic lattice. Each lattice site contains a planar rotor or a phase. In the continuum limit the Hamiltonian is still given by Eq. (II.2) except the integral over \mathbf{r} is now a d -dimensional integral and therefore the factor J is replaced by Ja^{2-d} . The average size of the projection of the rotors along, say, the x -direction in \mathbf{S} space, i.e. the magnetisation, is

$$\langle S_x \rangle = \langle \cos \theta(\mathbf{r}) \rangle \quad (\text{II.14})$$

$$= \langle \cos \theta(0) \rangle. \quad (\text{II.15})$$

Note that we might as well have chosen the y -direction. The model is isotropic and the x and the y directions are equivalent. When $\langle S_x \rangle \neq 0$ a preferred direction is singled out in the sense that on average \mathbf{S} points in the direction given by $\langle S_x \rangle$. In this case we say that the rotor field possesses order. In contrast if $\langle S_x \rangle = 0$ we also have $\langle S_y \rangle = 0$, since the model is isotropic. The zero projection comes about because the rotors circulate around and on average point equally much in all directions. So we say that the rotor field is disordered or does not possess any ordering.

First we neglect the singular vortex contributions (which is perfectly safe at low temperature) and Fourier transform the phase field

$$\theta(\mathbf{r}) = \int \frac{d\mathbf{k}}{(2\pi)^d} \hat{\theta}(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}} \quad (\text{II.16})$$

$$\theta(0) = \int \frac{d\mathbf{k}}{(2\pi)^d} \hat{\theta}(\mathbf{k}) \quad (\text{II.17})$$

$$\int d\mathbf{r} (\nabla \theta)^2 = \int \frac{d\mathbf{k}}{(2\pi)^d} k^2 \theta(\mathbf{k}) \hat{\theta}(-\mathbf{k}). \quad (\text{II.18})$$

These eqs. are substituted into the expression

$$\langle S_x \rangle = \frac{\int D[\theta] \cos(\theta(0)) e^{-\beta H}}{\int D[\theta] e^{-\beta H}} = \text{Re} \left(\frac{\int D[\theta] e^{-\beta H + i\theta(0)}}{Z} \right). \quad (\text{II.19})$$

After some algebra one obtains the following expression

$$\langle S_x \rangle = \exp \left(-\frac{T}{2Ja^{2-d}} \mathcal{S}_d \int_{\pi/L}^{\pi/a} dk k^{d-3} \right). \quad (\text{II.20})$$

The behaviour of $\langle S_x \rangle$ is controlled by the integral

$$I(L) = \int_{\pi/L}^{\pi/a} dk k^{d-3}. \quad (\text{II.21})$$

The behaviour of $I(L)$ strongly depends on the dimension d . For $d < 2$ we have $I(L) \sim L^{2-d} \rightarrow \infty$ as $L \rightarrow \infty$. Hence, $\langle S_x \rangle = 0$ in the limit of large systems for dimensions less than 2. For $d > 2$ we have that

$$I(L) \rightarrow A = \frac{1}{d-2} \left(\frac{\pi}{a} \right)^{d-2} \quad (\text{II.22})$$

and therefore

$$\langle S_x \rangle = \exp \left(-\frac{\mathcal{S}_d}{2Ja^{2-d}} AT \right) > 0. \quad (\text{II.23})$$

Finally for $d = 2$ the integral $I(L)$ is logarithmically divergent $I(L) = \ln(L/a)$ which is sufficient to force $\langle S_x \rangle$ to zero for any non-zero temperature.

We conclude that there is no ordered phase according to the behaviour of $\langle S_x \rangle$ at low temperature for $d \leq 2$. For $d < 2$ this means that there is no phase transition. The same was for a while thought to be the case for $d = 2$. Since the thermal motion included in the calculation of $\langle S_x \rangle$ is able to prevent a preferred direction and hence ensure $\langle S_x \rangle = 0$ for $T > 0$ including other types of excitations, such as vortices, can surely not make $\langle S_x \rangle$ different from zero. So it is safe to conclude that the rotors are unable to order along a common direction for any non-zero temperature and it was accordingly expected that a phase transition in the 2d YX-model was excluded. This conclusion was reached since in magnetic systems the average of the local magnetic moment, i.e. $\langle S_x \rangle$ is the order parameter and the phase transition takes place at the temperature where the order parameter changes from zero to a non-zero value. It turned out that by identifying the vortices as emergent collective excitations and by understanding their physical effects, a phase transition of a new kind was discovered in the 2d XY-model.

II.3 Vortex Unbinding

An indication of the importance of vortices as the temperature is increased can be obtain from the following simple and heuristic argument. We estimate the free energy

of a single vortex. The Helmholtz free energy is given by the difference between the energy and the entropy multiplied by the temperature $F = E - TS$. The energy is given by Eq. (II.11). We estimate the entropy from the number of places where we can position the vortex centre, namely on each of the $(L/a)^2$ plaquette of the square lattice, i.e., $S = k_B \ln(L^2/a^2)$. Accordingly, the free energy is given by

$$F = E_0 + (\pi J - 2k_B T) \ln(L/a). \quad (\text{II.24})$$

For $T < \pi J/2k_B$ the free energy will diverge to plus infinity as $L \rightarrow \infty$. At temperatures $T > \pi J/2k_B$ the system can lower its free energy by producing vortices: $F \rightarrow -\infty$ as $L \rightarrow \infty$. This simple heuristic argument points to the fact that the logarithmic dependence on system size of the energy of the vortex combines with the logarithmic dependence of the entropy to produce the subtleties of the vortex unbinding transition. Assume a different dependence of the energy on system size and one will either have thermal activation of vortices at all temperatures (in case $E_{vor} \rightarrow \text{const.} < \infty$) or vortices will not be activated at any temperature (in case $E_{vor} \sim (L/a)^b$ with $b > 0$). It is the logarithmic size dependence of the 2d vortex energy that allows the outcome of the competition between the entropy and the energy to change qualitatively at a certain finite temperature T_{KT} .

In reality it is not single vortices of the same sign that proliferate at a certain temperature. What happens is that the larger vortex pairs which are bound together for temperatures below T_{KT} unbind at T_{KT} . This is a collective effect that can be treated quantitatively by use of a special Renormalisation Group method design by Kosterlitz [41]. The vortex pairs induced as one approaches T_{KT} disturb the phase field so much that the effective value of the vortex binding term E_1 in the vortex pair free energy, that is Eq. (II.12) generalised to non-zero temperature, is driven to zero for large vortex separations. In the next section we shall see in detail how this happens.

II.3.1 The Spin Wave Stiffness

As our concern in this article is with emergent entities, we will now briefly discuss how a focus on, and an understanding of, the vortex degrees of freedom makes it possible to identify and describe the previously “hidden” phase transition in the 2d XY system.

The effect of the thermally activated vortex pairs is described by the temperature dependent *spin wave stiffness* ρ_s^R . This is an example of what Philip W Anderson calls a generalised rigidity[4]. The spin wave stiffness describes how much free energy it costs to apply a twist, or gradient, to the rotors (also called spins):

$$\theta(\mathbf{r}) = \theta_0(\mathbf{r}) + \mathbf{v}_{ex} \cdot \mathbf{r}, \quad (\text{II.25})$$

here $\theta_0(\mathbf{r})$ is allowed to vary according to the canonical ensemble. The increase in the free energy is given by

$$F(\mathbf{v}_{ex}) - F(0) = \frac{1}{2} V \rho_s^R v_{ex}^2. \quad (\text{II.26})$$

A number of comments concerning the notation are illuminating. The notation \mathbf{v}_{ex} for the gradient applied to the phase field $\theta(\mathbf{r})$ has its origin in the fact that the same physics, as we describe here, applies to superfluid films and superconducting films. In these cases the field $\theta(\mathbf{r})$ is the phase of the complex order parameter, the wave function of the superfluid. Being the phase of a quantum mechanical wave function the gradient of $\theta(\mathbf{r})$ is related to a probability current and thereby to the velocity field of the superfluid. The notation ρ_s^R is meant to remind one that this phase rigidity is determined by the **density** of superfluid in the case of a superfluid or a superconductor. The superscript R in ρ_s^R indicates that thermal excitations renormalise the quantity. It follows immediately from the Hamiltonian in Eq. (II.2) that at zero temperature $\rho_s^R = J = \rho_s$. The spin wave stiffness is similar to the shear constant of a material. The shear constant determines how the (free) energy increase when a shear deformation is imposed. As temperature is increased the shear constant decreases and drops abruptly to zero when the solid melts into a liquid.

To obtain ρ_s^R one calculates the left hand side of Eq. (II.26). Details can be found in the wonderful book by Chaikin and Lubensky[12]. The phase field is split into two parts

$$\theta_0(\mathbf{r}) = \theta_s(\mathbf{r}) + \theta_v(\mathbf{r}), \quad (\text{II.27})$$

where the first term describes smooth spin waves and the second term contains the singular vortex contribution. The free energy is obtained from $F = k_B T \ln Z$ and the partition function is given by Eq. (II.3). To calculate Z introduce Fourier transforms of the phase field. After quite a bit of algebra one arrives at the following simple expression

$$\rho_s^R = \rho_s - \frac{1}{2} \frac{\rho_s^2}{T} \lim_{k \rightarrow 0} \frac{\langle \hat{n}(\mathbf{k}) \hat{n}(-\mathbf{k}) \rangle}{k^2}, \quad (\text{II.28})$$

which expresses the renormalised stiffness in terms of the correlation function of the Fourier transform of the vortex density function

$$n(\mathbf{r}) = \sum_{\alpha} n_{\alpha} \delta(\mathbf{r} - \mathbf{r}_{\alpha}), \quad (\text{II.29})$$

for a collection of vortices of charge n_{α} (see Eq. (II.6)) with centres located at positions \mathbf{r}_{α} . The vortices are now described entirely by their position exactly like if they were ordinary particles. So what started out as a complex configuration in the field of rotors

is now possible to treat as point particles. The effect of the extended disturbance of the rotor field is taken care of by the interaction energy between two vortices. The thermodynamic average in Eq. (II.28) is over the canonical ensemble with no twist imposed, hence the subscript 0. Eq. (II.28) can be used to determine how the spin wave stiffness behave at large distances as a function of temperature. We will discuss how in the next section.

II.3.2 The KT transition

Let us first summarise the phenomenology of the Kosterlitz-Thouless transition. As the temperature is increased more and more vortex pairs are thermally activated. This makes ρ_s^R decrease, see Eq. (II.28). This corresponds to a decrease in the increment of the free energy induced by a certain twist \mathbf{v}_{ex} . We can understand the effect from the fact that the phase field $\theta(\mathbf{r})$ becomes more and more distorted as the temperature is increased, hence the extra perturbation caused by \mathbf{v}_{ex} becomes relatively less important. Quantitatively one finds

$$\rho_s^R = \begin{cases} \rho_s^R(T_{KT}^-)[1 + const.(T_{KT} - T)^{1/2}] & \text{for } T < T_{KT} \\ 0 & \text{for } T > T_{KT}. \end{cases} \quad (\text{II.30})$$

Here, T_{KT} is the Kosterlitz-Thouless temperature at which vortex pairs unbind. The value of T_{KT} differs from one system to another. In the 2d XY-model $T_{KT}/J \simeq 0.893 \pm 0.002$ [52]. The remarkable thing is, as we shall see below, that the ratio

$$\rho_s^R(T_{KT}^-)/T_{KT} = 2/\pi \quad (\text{II.31})$$

is universal for all systems that undergoes a KT-transition. Since $\rho_s^R(T_{KT}^+) = 0$ Eq. (II.31) is referred to as the universal jump. The correlation length $\xi(T)$ behaves in a very unusual way as one approaches T_{KT} from above. We are used to a relatively slow algebraic divergence of the correlation length as the critical temperature is approached. For the KT-transition the divergence is, however, much faster

$$\xi(T) \sim \exp\left(\frac{const.}{(T - T_{KT})^{1/2}}\right) \text{ for } T > T_{KT}. \quad (\text{II.32})$$

Can we in a simple way understand this exponential divergence? Yes, we can. The phase field is significantly distorted by unbound vortices, since these vortices are not screened by a nearby anti-vortex. I.e. the phases $\theta(\mathbf{r})$ can remain correlated over distances shorter than the typical distance $\langle D \rangle = 1/\sqrt{n_{ub}}$ between unbound vortices of density n_{ub} [37]. Or in other words, we expect the correlation length $\xi \sim D$. The

vortices are thermally induced and therefore their density is expected to depend on the temperature through a Boltzmann factor $\exp(-E_{\text{vor}}/T)$. The situation described here is exactly what happens in the one dimensional so-called ϕ^4 model. This model supports thermally activated solitons. The correlation length is set by the inverse of the soliton density and diverges exponentially as the temperature goes to zero [36]. The same thing, in a slightly simpler version, also happens in the one dimensional Ising model. This argument can indicate the cause of the exponential dependence of ξ . But it is no more than an indication since the exponential dependence in Eq. (II.32) is significantly different from a simple Boltzmann factor. This difference is due to corrective renormalization effects.

Continuous phase transitions are accompanied by divergences in thermodynamic quantities caused by the divergence of the correlation length as the critical temperature T_c is approached. The singular part of the free energy density f can be estimated as the amount of thermal energy T_c within a correlated volume ξ^d , which gives $f \sim T_c/\xi^d$. The specific heat c_V is given by the second derivative of the free energy $c_V = -T\partial^2 f/\partial T^2 \sim \partial^2 \xi^{-d} \partial T^2$. For the KT-transition the exponential divergence of $\xi(T)$ in Eq. (II.32) is so rapid and occur over such a narrow temperature range that the divergence in c_V cannot be resolved in simulations or in experiment. This is another reason why the vortex unbinding transition remained unnoticed for so long. It doesn't leave any dramatic signature in the thermodynamic quantities. However, as mentioned above, the macroscopic rigidity clearly changes at the transition.

II.4 The Vortex Unbinding Transition in Other Systems

We have above mentioned that not only the XY-model exhibit the Kosterlitz-Thouless vortex unbinding transition[46]. Any two dimensional system that supports thermally induced “charges” or topological defects that interact logarithmically will undergo this transition. The $U(1)$ symmetry of the phase field $\theta(\mathbf{r})$ of the XY-model is also present in the Ginzburg-Landau free energy of superfluids and of superconductors. The topological excitations in the case of a superfluid consist of vortices in the flow of the superfluid. Vortices like those observed when one empties a bath tub. In thin superfluid helium film such vortices destruct the superfluid phase with increasing temperature according to the scenario of the KT-transition[1, 2].

The situation is slightly more complicated in superconductors. Because the superfluid in this case is charged (the superconducting pairs of electrons), screening effects play a role[73, 46, 12]. However, for thin superconducting films of thickness δ the

effective screening length is given by $\lambda_{eff} = \lambda^2/\delta$, which can easily become a macroscopic length. In this case the loss of superconductivity is caused by the unbinding of vortex pairs according to the KT-transition. The broken pairs can move freely when they respond to an applied electric current. As they move they cause phase slips in the superconducting order parameter. These phase slips induce a voltage drop according to the Josephson relation. The superconductor is now unable to support an electric current without a voltage drop, i.e. it is not a superconductor any longer[39, 34, 18, 6]

Dislocations in two dimensional crystals interact through the strain field. Two edge dislocations of opposite sign correspond to an extra row of atoms inserted along the line connecting the location of the two dislocation cores. The extra line of atoms produces strain and leads to an increase in the energy which is logarithmic in the separation between the two dislocations. Thus the situation is very similar to the one encountered in the XY-model. When the dislocations unbind, free dislocations are produced. A shear applied to the system can now be accommodated by the mobile dislocations without an increase in the (free) energy. I.e., the shear constant has dropped to zero and the system is melted. The 2d melting theory of Kosterlitz-Thouless-Halperin-Nelson-Young predicts that melting occur in two stages. At the first stage dislocations unbind and make the shear constant drop to zero. The dislocations are topological defects, their effect on the order of the lattice are, however, not very dramatic. Before the unbinding of dislocations, the translational and the orientational order of the lattice are both described by correlation functions that depend algebraically on distance. When the dislocations unbind the translational correlation function becomes exponential but the orientational correlations remain algebraic. At a somewhat higher temperature topological defects called disclinations unbind with the effect that the orientational order becomes exponential. Details can be found in Chaikin and Lubensky [12].

There are many other cases where the logarithmic vortex interaction and the KT-transition play a role. For instance, the shape of surfaces in three dimensions may undergo a transition from smooth to rough[5]. Assume that the surface energy of the two dimensional surface is proportional to the area of the surface. And assume that the surface is defined in terms of its height $h(x, y)$ above the xy -plane, i.e. no overhangs. In other words the points on the surface have the coordinates $(x, y, h(x, y))$. The Hamiltonian for the surface is then

$$H = \sigma \int dx \int dy \sqrt{1 + (\nabla h)^2}. \quad (11.33)$$

Here σ is a measure of the surface tension. If the height only varies slowly as function of (x, y) we can assume $|\nabla h| \ll 1$ and then expand the square root. In this approxi-

mation the Hamiltonian in Eq. (II.33) can be written as

$$H = \sigma L^2 + \frac{1}{2} \int dx \int dy (\nabla h)^2, \quad (\text{II.34})$$

(L is the linear size of the system in the xy -plane) which is equivalent to Eq. (II.9), and we expect the same physical phenomenology to apply to the surface as we found for the XY -model.

Chapter 12

Forecasting

We include the paper:

D. Piovani, J. Grujic and H.J. Jensen, Linear Stability Theory as an early warning sign for transitions in High Dimensional Complex Systems, *J. Phys. A: Math. Theor.* **49**, 295102 (2016)

which analyse in detail a new approach to the monitoring and forecasting of the onset of transitions in high dimensional complex systems by application to the Tangled Nature model of evolutionary ecology and high dimensional replicator systems with a stochastic element. This forecasting technique is relevant to situations where the considered system evolve and agents of a new type may appear as a result of the evolutionary dynamics.

Linear stability theory as an early warning sign for transitions in high dimensional complex systems

Duccio Piovani^{1,2}, Jelena Grujić^{1,3,4} and
Henrik Jeldtoft Jensen^{1,5}

¹ Centre for Complexity Science and Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

² Centre for Advanced Spatial Analysis, UCL London, 90, Tottenham Court Road, London W1T 4TJ, UK

³ AI-LAB Computer Science Department, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

⁴ MLG, Département d’Informatique, Université Libre de Bruxelles, Boulevard du Triomphe—CP 212 1050 Brussels, Belgium

E-mail: duccio.piovani@gmail.com, jelenagr@gmail.com and h.jensen@imperial.ac.uk

Received 22 February 2016, revised 20 April 2016

Accepted for publication 6 May 2016

Published 10 June 2016



Abstract

We analyse in detail a new approach to the monitoring and forecasting of the onset of transitions in high dimensional complex systems by application to the Tangled Nature model of evolutionary ecology and high dimensional replicator systems with a stochastic element. A high dimensional stability matrix is derived in the mean field approximation to the stochastic dynamics. This allows us to determine the stability spectrum about the observed quasi-stable configurations. From overlap of the instantaneous configuration vector of the full stochastic system with the eigenvectors of the unstable directions of the deterministic mean field approximation, we are able to construct a good early-warning indicator of the transitions occurring intermittently.

Keywords: forecasting critical transitions, adaptive systems, evolutionary dynamics, replicator system

(Some figures may appear in colour only in the online journal)

⁵ Author to whom any correspondence should be addressed.

1. Introduction

High dimensional complex systems, both physical and biological, exhibit intermittent dynamics, consisting of stretches of relatively little change interrupted by often sudden and dramatic transitions to a new meta-stable configuration [1]. Such transitions can have crucial consequences when they occur in, say, ecosystems or financial markets and it is therefore important to develop methods that are able to identify precursors, warning signals and ideally techniques to forecast the transitions before they take place. We will expect that the mechanisms behind the rapid rearrangement may be different in different systems.

The literature in this field is very extensive and it is difficult to produce a comprehensive review. Here we relate to the literature most relevant to our paper. An often used approach to forecasting in complex systems was introduced by Scheffer and collaborators [2, 3], who suggest that critical slowing down and enhanced fluctuations can be used as a precursor of approaching systemic change. The method can work when the high dimensional dynamics of the complex systems can be captured by some few macroscopic collective degrees of freedom. It is expected to be of particular relevance when there is a slow change in some external parameter which drives the system towards the bifurcation point. In this case transitions are called *critical transitions* and the mathematical framework used comes from dynamical systems theory [4–6]. This method is pertinent to systems that are dynamically effectively low dimensional, in which the transition takes the form of a bifurcation captured by a robust macroscopic variable, which emerges from the micro dynamics. This approach has been applied for many years in many different fields, in climate change [7, 8], population dynamics [9], ecosystems [10] and very recently in financial markets [11] to mention just some examples. Furthermore, exploiting the same mathematical framework *flickering* between two stationary states has been found to be a useful indicator [12, 13].

More recently in [14] the authors describe transitions in a different way. In this new interpretation transitions are named *saddle-escape transitions* and metastable states are interpreted as high dimensional saddle points. Transitions are not induced as a result of a change in the external parameters, as is the case in the bifurcation interpretation, but happen due to a rare perturbation which pushes the system towards an unstable direction. An early warning sign is then captured by inferring the value of the largest eigenvalue of the Jacobian through the log difference of a conveniently chosen macroscopic time series.

Here we give the same interpretation of metastable states, though we develop a different approach. As we suggested recently in [15], transitions are induced by intrinsic fluctuations at the level of the individual components which propagate to the macroscopic systemic level and thereby trigger a change in the overall configuration. Our approach is relevant to systems in which the available configuration space evolves as a consequence of the dynamics. One may think of a new and more virulent virus being created through a mutation of an existing strain (e.g. the SARS virus in 2003), or a new economic agent arriving in the market (e.g. the dot-com bubble in 1997-2000). Contrary to [14], to build our indicator we do not make use of a rigid macroscopic variable to monitor the system; instead we analyse the interactions between the single microscopic components.

In the current paper we elaborate the method we presented in [15], we discuss the mathematical details both of the mean field approximation and of a linear stability analysis (LSA), and explore in greater detail its forecasting power. As a first test case we consider the Tangled Nature (TaNa) model of evolutionary ecology [16], which has had considerable success in reproducing both macro-evolutionary aspects such as the intermittent mode of extinctions [17] and ecological aspects such as species abundance distributions [18] and species area laws [19]. A much more succinct and schematic discussion was given in [15].

Furthermore, as a new test case, we present results for transitions in a model with a very different type of dynamics, namely a high dimensional replicator with a stochastic element of mutation [20, 21]. This model is based on the replicator–mutator equation which has wide application in many different fields, such as population genetics [22], evolutionary game theory [23], language evolution [24], etc. Furthermore, it is related to a few different models with wide applications on their own. First, it contains the mutation element of the quasispecies equation [25, 26] often used to model the spread of strongly mutating viruses like HIV and Hepatitis C [27] and the frequency dependent element of the replicator equation [28] used in evolutionary game theory [29]. The Lotka–Volterra equation [30], otherwise known as predator–prey model, which was originally developed to describe dynamics in ecological systems and later was widely applied in economics [31], has been shown to be equivalent to the replicator equation [32]. The Price equation [33], which Hamilton used in his work on kin selection [34] and was later used to describe numerous biological systems, is also equivalent to the replicator equation, and the expanded Price equation is equivalent to the replicator–mutation equation [35]. Given the broad relevance of the replicator–mutation equation (population dynamics, virus spreading, game theory, financial dynamics, social dynamics etc.), success in forecasting transitions in this model may indicate that our method might be relevant to a range of situations. We begin by demonstrating that the high dimensional replicator system with mutations exhibits intermittent behaviour. Without mutations, the replicator equation will typically not exhibit intermittent dynamics.

Despite their different general mechanisms, the TaNa model and the replicator–mutator system share similar properties. Their stochastic dynamics are characterised by a huge number of metastable states. When the system randomly falls into one of these it enters a quiescent period of little change. Eventually the intrinsic stochastic fluctuations lead to the occupancy of hitherto empty parts of the configuration space, which may serve as a random kick able to drive the system away from the metastable configuration and towards the chaotic regime, where the system undergoes a high dimensional adaptive walk searching for another metastable point. The two systems studied here do not exhibit the characteristic bifurcation captured by some rigid macrovariable, nor can the transitions be forecasted by inferring the largest eigenvalue. They cannot be forecasted through the observation of systemic properties but require the analysis of microscopic interactions.

Nevertheless, through a mean field description of the stochastic dynamics we can infer the Jacobian and interpret the metastable states as high dimensional saddle points formed of a vast majority of stable directions and typically a few unstable ones. If the mean field description was accurate, we would be able to make deterministic predictions. This is not the case, but we demonstrate that we are able, in both models, to understand which kind of intrinsic stochastic fluctuation will be able to push the system out of its stable configuration.

As said in this paper we start by elaborating on the procedure presented in [15], checking the performance of an alarm threshold built on the stability indicator. This procedure is to be considered a starting point; its weakness is the need of full information on the system (one needs to know the structure of the whole configuration space). To overcome this problem we have developed a new methodology, described at the end of the paper, where we have reduced the amount of dynamical detail needed to produce forecasts.

The remaining of the paper is structured as follows: in section 2 we describe the procedure which combines observational data (in our case from simulations) with an LSA of the mean field dynamics, in section 3 we will go through the details of the models used as test cases and their mean field description. In section 4 we analyse the results coming from the forecasting procedure first introduced in [15] applied to both models. In section 5A, to study the robustness of the method, we introduce a level of error in the mean field interaction

matrix. Finally in section 5B we introduce and develop the methodology to make the method more applicable to real world problems.

In order to facilitate collaboration we have uploaded all the codes (C/C \times \times) necessary to produce the results of the paper online (both the models and the forecasting procedure). The interested reader can find and download them from H J Jensen's home page <https://wwwf.imperial.ac.uk/~hjens/>.

2. Linear stability analysis application

In this section we give a general outline of our approach. We will then describe the application to the two models in detail in the following two sections. The first step is to establish a mean field approximation of the stochastic dynamics in order to obtain a set of deterministic equations. In order to do so we define the state vector $\mathbf{n}(t) = (n_1(t), \dots, n_d(t))$, where $n_i(t)$ represents the occupation of each node (species or strategy). The mean field time evolution is of the form

$$\mathbf{n}(t+1) - \mathbf{n}(t) = \mathbb{T}(\mathbf{n}(t)) \cdot \mathbf{n}(t) \quad (1)$$

where the matrix \mathbb{T} is the mean field evolution matrix, which will contain contributions from the following processes: death, reproduction and mutation. Obviously in this framework the fixed point configurations \mathbf{n}^* are given as solutions of

$$\mathbb{T}(\mathbf{n}^*) \cdot \mathbf{n}^* = 0. \quad (2)$$

Because of the high dimensionality of the type of systems we have in mind, equation (2) will typically not be solvable analytically but can be numerically approximated observing the stochastic dynamics. We now perform a linear stability analysis about \mathbf{n}^* by introducing a small perturbation $\mathbf{n}(t) = \mathbf{n}^* + \delta\mathbf{n}(t)$. By substituting the perturbed vector in equation (1) and expanding the right hand side to first order in $\delta\mathbf{n}(t)$ we get

$$\begin{aligned} \delta\mathbf{n}(t+1) - \delta\mathbf{n}(t) &\simeq (\mathbb{T}(\mathbf{n}^*) + \partial_{\mathbf{n}}\mathbb{T}(\mathbf{n}^*) \cdot \mathbf{n}^*)\delta\mathbf{n}(t) \\ &= \mathbb{M}(\mathbf{n}^*) \cdot \delta\mathbf{n}(t) \end{aligned} \quad (3)$$

where we have used equation (2). Here the matrix

$$\mathbb{M}(\mathbf{n}^*) = (\mathbb{T}(\mathbf{n}^*) + \partial_{\mathbf{n}}\mathbb{T}(\mathbf{n}^*)\mathbf{n}^*) \quad (4)$$

is the *Jacobian* of the system, or the *stability matrix*. Now exploiting the results of the LSA, we know that the eigenvectors or *generalised* eigenvectors (in the case of a non diagonalizable Jacobian) \mathbf{e}_+ associated with λ with $Re(\lambda) > 0$ indicate unstable directions. These can be identified with *dangerous* components n_i of the configuration vector towards which the eigenvectors point. What this means is that if the stochastic fluctuations (mutations) bring the system close to these unstable directions, by activating the *dangerous* components, the system will suffer a repulsive force that will push it away from the fixed (saddle) point \mathbf{n}^* . In other words the *activation* of one of these components corresponds to a perturbation parallel to an unstable direction of the saddle fixed point \mathbf{n}^* . This implies that a sudden growth of these components would indicate the arrival of a transition.

This observation allows us to identify a stability indicator, whose non-zero values are early warning signalling of an approaching transition caused by the system leaving the vicinity of a current fixed point. The details of this indicator will depend on the specific case we are dealing with but will be based on the same general idea. In the following sections we will present the two test case models analysing their basic mechanisms and results, and developing our mean-field stability indicator in both cases.

3. The models

3.1. A. The Tangled Nature model

In the TaNa, an agent is represented by a sequence of binary variables with fixed length L , denoted as $\mathbf{S}^a = (S_1^a, \dots, S_L^a)$, where $S_i^a = \pm 1$. Thus, there are 2^L different sequences, each one represented by a vector in the genotype space: $\mathcal{S} = \{-1, 1\}^L$. In a simplistic picture, each of these sequences represents a genome uniquely determining the phenotype of all individuals of this genotype. We denote by $n(\mathbf{S}^a, t)$ the number of individuals of type \mathbf{S}^a at time t and the total population is $N(t) = \sum_{a=1}^{2^L} n(\mathbf{S}^a, t)$. We define the distance between different genomes \mathbf{S}^a and \mathbf{S}^b as the Hamming distance: $d_{ab} = \frac{1}{2L} \sum_{i=1}^L |S_i^a - S_i^b|$. A time step is defined as a succession of one annihilation and of one reproduction attempt. During the killing attempt, an individual is chosen randomly from the population and killed with a probability p_{kill} constant in time and independent of the type. During the reproduction process, a different randomly chosen individual \mathbf{S}^a successfully reproduces with probability $p_{\text{off}}(\mathbf{S}^a, t) = \frac{\exp(H(\mathbf{S}^a, t))}{1 + \exp(H(\mathbf{S}^a, t))}$, which depends on the occupancy distribution of all the types at time t via the weight function:

$$H(\mathbf{S}^a, t) = \frac{k}{N(t)} \sum_{\mathbf{S}^b \in \mathcal{S}} \mathbf{J}(\mathbf{S}^a, \mathbf{S}^b) n(\mathbf{S}^b, t) - \mu N(t). \quad (5)$$

In equation (5), the first term couples the agent \mathbf{S}^a to one of type \mathbf{S}^b by introducing the interaction strength $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$, whose values are randomly distributed in the interval $[-1, +1]$. For simplicity, and to emphasise interactions, we here assume: $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^a) = 0$. The parameter k scales the interaction strengths and μ can be thought of as the carrying capacity of the environment. An increase (decrease) in μ corresponds to harsher (more favourable) external conditions. The reproduction is asexual: the reproducing agent is removed from the population and substituted by two copies \mathbf{S}_1^a and \mathbf{S}_2^a , which are subject to mutations. A single mutation changes the sign of one of the genes: $S_i^a \rightarrow -S_i^a$ with probability p_{mut} . Similarly to a Monte Carlo sweep in statistical mechanics, the unit of time of our simulations is a *generation* consisting of $N(t)/p_{\text{kill}}$ time steps, i.e. the average time needed to kill all the individuals at time t . These microscopic rules generate intermittent macro dynamics. The system is persistently switching between two different modes: the meta-stable states (denoted quasi-evolutionary stable strategies (QESS) and the transitions separating them. The QESS states are characterised by small amplitude fluctuations of $N(t)$ and stable patterns of occupancies of the types (figure 1, left and right panel respectively). However, these states are not perfectly stable and configurational fluctuations may trigger an abrupt transition to a different QESS state. The transitions consist of collective adaptive random walks in the configuration space while searching for a new metastable configuration and are related to high amplitude fluctuations of $N(t)$. All the results we will present for this model have been obtained by fixing the parameters to $L = 8$, $p^{\text{mut}} = 0.2$, $p^{\text{kill}} = 0.4$, $K = 40$ and $\mu = 0.07$ and have been chosen for computational reasons. Furthermore, one can see from the occupancy plot in the right panel how a QESS configuration only occupies a small portion of the total available genome space (blue dots). With this parameter set, typically 20–50 nodes of the genome space are active, 5–10 of which are being heavily occupied (wild types) out of the 256 available.

Should be
 $p_{\text{mut}} = 0.02$
 $\mu = 0.007$

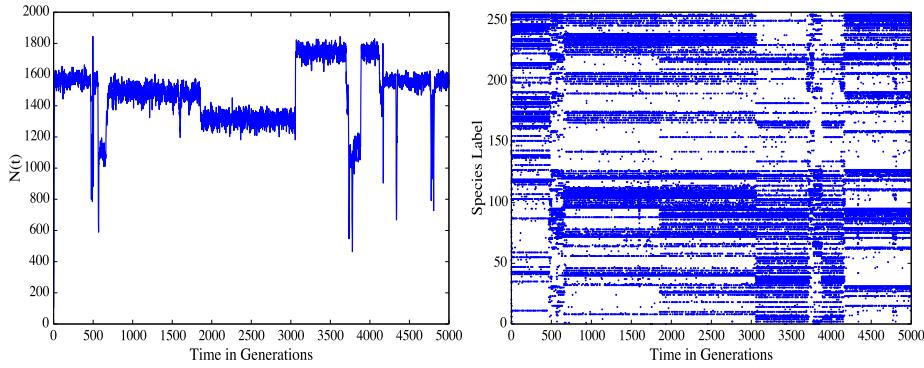


Figure 1. Left panel: total population as a function of time (in generations) for a single realization of the TaNa. The punctuated dynamics are clearly visible: quasi-stable periods alternate with brief periods of hectic transitions, during which $N(t)$ exhibits large amplitude fluctuations. Right panel: occupancy distribution of the types. The genotypes are labelled arbitrarily and a dot indicates a type which is occupied at the time t . These figures are obtained with parameters $L = 8$,
Should be $p^{\text{mut}}=0.02$ $p^{\text{mut}} = 0.2$, $p^{\text{kill}} = 0.4$, $K = 40$ and $\mu = 0.007$.

3.2. Mean field description

In the TaNa model there are multiple sources of stochasticity, namely reproduction, mutations and deaths. Following the procedure outlined above we average over these sources in order to derive a deterministic mean field equation. At each time step with probability p_{kill} a randomly chosen individual is removed from the system, which implies that the occupation number of the species it belongs to decreases by one ($\Delta n_i = -1$). Given that the probability of choosing an individual belonging to the i th species is $\rho_i = \frac{n_i}{N}$, the probability that an individual of type i is removed is given by

$$\rho_i \cdot p_{\text{kill}} \cdot (-1). \quad (6)$$

The reproduction term is slightly more complicated because mutations need to be treated with some care. A randomly chosen individual is selected for asexual reproduction, which means it is removed from the system while creating two new individuals of the same species. Offsprings can both mutate ($\Delta n_i = -1$), only one can mutate ($\Delta n = 0$), or none mutate ($\Delta n = +1$). Keeping in mind that the probability of reproducing is given by p_i^{off} the average contribution from reproduction of type i including mutations is

$$\rho_i \cdot p_i^{\text{off}}(t)[2p_o - 1] = \alpha \cdot \rho_i \cdot p_i^{\text{off}}(t) \quad (7)$$

where $p_o = (1 - p^{\text{mut}})^L$ is the probability of no mutations and $\alpha = (2p_o - 1)$ is a constant. The third term we have to consider is the *backflow effect*, which describes the event of being populated by mutations occurring during the reproduction happening elsewhere. This term has the form

$$\sum_j \rho_j(t) p_j^{\text{off}} p_{j \rightarrow i}^{\text{mut}}. \quad (8)$$

Type j will have to mutate a number of genes corresponding to its hamming distance d_{ij} between j and i in order to increase n_i . This will happen with probability

$$p_{i \rightarrow j}^{\text{mut}} = p_{\text{mut}}^{d_{ij}} \cdot (1 - p_{\text{mut}})^{L-d_{ij}}. \quad (9)$$

Putting together all these effects we find the expression for equation (1) for this model to be

$$\begin{aligned} & n_i(t+1) - n_i(t) \\ &= \frac{1}{N} \sum_j \{ (p_j^{\text{off}}(t)(2p_o - 1) - p^{\text{kill}}) \cdot \delta_{ij} + p_j^{\text{off}} \cdot p_{j \rightarrow i}^{\text{mut}} \cdot (1 - \delta_{ij}) \} n_j(t) \end{aligned} \quad (10)$$

where

$$T_{ij} = (p_j^{\text{off}}(t)(2p_o - 1) - p^{\text{kill}}) \cdot \delta_{ij} + p_j^{\text{off}} \cdot p_{j \rightarrow i}^{\text{mut}} \cdot (1 - \delta_{ij}) \quad (11)$$

is the mean-field evolution matrix of the system. By substituting equation (11) into equation (4) we get the specific form of the stability matrix for the TaNa model

$$\begin{aligned} \mathbb{M}_{ij} &= (\alpha p_j^{\text{off}} - p^{\text{kill}}) \delta_{ij} + 2(1 - \delta_{ij}) p_j^{\text{off}} p_{j \rightarrow i}^{\text{mut}} \\ &+ \sum_k [\alpha \delta_{ik} + (1 - \delta_{ik}) \cdot p_{k \rightarrow i}^{\text{mut}}] \frac{\partial p_k^{\text{off}}}{\partial n_j} n_k^*. \end{aligned} \quad (12)$$

This is the mean field matrix we will use for our linear stability analysis of the stochastic fixed points.

3.3. B. The replicator model with mutations

The replicator equation [28] was introduced in evolutionary game theory in order to capture the frequency-dependent nature of the evolution process. Namely, in this model the fitness, or the evolutionary success, of a strategy depends on the frequency of the other strategies in the system. By combining the replicator equation with the quasispecies equation we obtain the replicator–mutation equation, where, apart from the frequency dependence, we also allow for new strategies to enter the system through mutations. As mentioned in the introduction, this model is used to describe numerous high dimensional socio-economic or biological systems.

We are interested in the limit of many strategies. Players may leave the system (say, go bankrupt or extinct) or may change their strategy (mutate). This means that the number of players choosing a given strategy and the number of available strategies are in constant evolution. This version of the replicator dynamics set-up was studied by Tokita and Yasutomi in [21]. The authors focused on the emerging network properties. Here we continue this study but with an emphasis on the intermittent nature of the macro-dynamics.

For this model the configuration vector $\mathbf{n}(t)$ contains the *relative frequencies* of all the allowed d different strategies, so the components $n_i(t) \in [0, 1]$ for all $i = 1, 2, \dots, d$. Not all strategies need to be *active* at a given moment, i.e. we can have $n_i(t) = 0$ for some strategy i . We start the simulations by generating the $d \times d$ payoff matrix J of the game that will tell us the payoffs of every pairwise combination. In the same way as the TaNa model above, the matrix J is a random and fixed interaction network on top of which the replicator dynamics will evolve. Each strategy distinguishes itself from the others in its payoffs or interactions with the rest of the strategy space.

In this chapter we used the same type of uncorrelated interaction matrix as used in the study above of TaNa model. The dimension of the matrix is large, namely $d \in (10^2, 10^4)$. The qualitative aspects of the behaviour remain the same for other types of payoff matrices. We found that matrices with payoffs uniformly distributed on the interval $(-1, 1)$ or on the set $\{0, 1\}$ exhibit the same behaviour as matrices of the form used for the TaNa model.

However, if the payoffs are drawn from a power law distribution with no second moment, the dynamics become different and the intermittent behaviour is no longer so distinct.

In the initial configuration, $N_o < d$ strategies start with the same frequency $n_i = \frac{1}{N_o}$. All the other possible strategies are non active, i.e. the corresponding $d - N_o$ components in $\mathbf{n}(0)$ are $n_i(0) = 0$. The empty strategies can only become populated by one of the *active* strategies mutating into them. Once this happens their frequency will evolve according to the replicator equation in which these newly occupied strategies interact with the active strategies which they are linked to through the matrix J .

A time step of the replicator dynamics consists of calculating the *fitness*, $h_i(t) = \sum_j J_{ij} n_j(t)$ of each active strategy and comparing it with the average fitness $\bar{h}(t) = \sum_{ij} J_{ij} n_i(t) n_j(t)$, exactly as expected in a replicator dynamics. Each frequency is then updated according to

$$n_i(t+1) = n_i(t) + \left(\sum_j J_{ij} n_j(t) - \sum_{kj} J_{kj} n_k(t) n_j(t) \right) n_i(t). \quad (13)$$

The stochastic element, of the otherwise deterministic dynamics, consists in the following updates. With probability p^{mut} each strategy mutates into another, this is done by transferring a fraction α_{mut} of the frequency from the considered strategy to another strategy. The label of the latter strategy is chosen in the vicinity of the first by use of a normal distribution $N(i, \Delta)$ centred on label $i \in \{1, 2, \dots, d\}$ with variance Δ with periodic boundary conditions, i.e. label $d + 1 = 1$. The closer the labels of two strategies are, the more likely it is for one to mutate into the other.

It should be noted that as long as the payoff matrix is random and uncorrelated in its indices, strategies which have a small difference between their indices are not necessarily similar since the two strategies may interact with the other strategies in completely different ways. The Δ parameter is introduced to control the level of stochasticity in the system. The higher the values of Δ , the greater the range of nodes j that can be activated from a mutation happening in i .

When the frequency of a strategy i goes below a preset extinction threshold $n_i(t) < n^{\text{ext}}$, the strategy is considered extinct and its frequency is set to zero $n_i(t+1) = 0$. Right after an extinction event the system is immediately renormalised in order to maintain the condition $\sum_i n_i(t) = 1$.

The dynamics at the systemic level are captured by the time evolution of the occupancy vector $\mathbf{n}(t)$ and shown in figure 2, where we present the *occupancy* plot (left panel) and the evolution of the frequencies of the single strategies (right panel).

All the results for this model have been obtained with the same parameter set, namely: $d = 256$, $n^{\text{ext}} = 0.001$, $\alpha_{\text{mut}} = 0.01$, $p^{\text{mut}} = 0.2$ and $\Delta = 15$, which once again have been chosen for reasons of computational performance. In this model the QESS are typically characterised by two strongly occupied strategies which are surrounded by seven to eight *cloud* strategies which are populated by mutations and quickly die out. So once again the stable configurations occupied only a small part of the entire strategy space.

3.4. Mean field description

The random mutations are the only source of stochasticity in the model's dynamics. To account for these stochastic events one has to consider the possibility that a strategy loses part of its frequency by mutating into other strategies or gains frequency as a result of mutations happening elsewhere. As a result a given strategy may lose a fraction of players

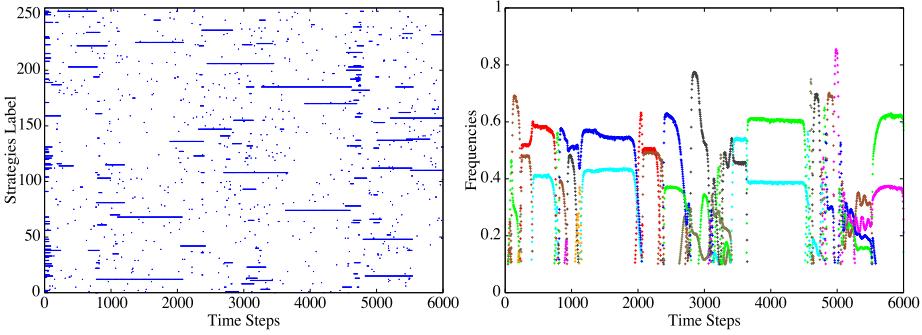


Figure 2. Left panel: occupancy distribution of the types. The genotypes are labelled arbitrarily and a dot indicates a type which is occupied at the time t . The punctuated dynamics is clearly visible: quasi-stable periods alternate with brief periods of hectic transitions. Right panel: the frequencies of the strategies. Each colour belongs to a different strategy. Once again the transitions from one meta stable configuration (approximate fixed point) to another is clear.

α_{mut} , which happens with probability p_{mut} or gain $\alpha_{\text{mut}} \cdot n_j(t+1)$ which happens with probability $p_{\text{mut}} \sum_{j \in N_a} p_{j \rightarrow i}$, where N_a is the number of active strategies and

$$p_{j \rightarrow i} = \frac{e^{-|i-j|^2}}{\sqrt{2\pi\Delta^2}} \quad (14)$$

is the probability of i mutating into j (or viceversa). This second effect describes the probability of being populated by a mutation. We therefore get the mean field description as

$$\begin{aligned} n_i(t+1) &\simeq n_i(t) + \left(\sum_j J_{ij} n_j(t) - \sum_{jk} J_{ik} n_i(t) n_k(t) \right) \cdot n_i(t) \\ &+ p_{\text{mut}} \alpha_{\text{mut}} \left(\sum_j n_j(t) p_{j \rightarrow i} - n_i(t) \right) \end{aligned} \quad (15)$$

which can be expressed, in compact form as

$$n_i(t+1) - n_i(t) \simeq \sum_j \mathbb{T}_{ij} n_j(t) \quad (16)$$

where

$$\begin{aligned} \mathbb{T}_{ij} &= \left(\sum_j J_{ij} n_j(t) - \sum_{jk} J_{ik} n_k(t) n_j(t) - p_{\text{mut}} \alpha_{\text{mut}} \right) \cdot \delta_{ij} \\ &+ p_{\text{mut}} \alpha_{\text{mut}} p_{j \rightarrow i} \cdot (1 - \delta_{ij}). \end{aligned} \quad (17)$$

The stability matrix is obtained by substituting equation (17) in equation (4)

$$\mathbb{M}_{ij} = \mathbb{T}_{ij}(\mathbf{n}^*) + \left[J_{ij} - \sum_k (J_{ik} + J_{ki}) \mathbf{n}_k^* \right] \mathbf{n}_i^*. \quad (18)$$

4. Procedure and results

We described in the previous sections how the dynamics of the two models consists in intermittent swift transitions between quasi-metastable configurations. As already stated it is not possible to analytically solve equation (2) but we can approximate the fixed points of the mean field dynamics by local time averages over successive configurations in the quasi-stable phases of the full stochastic dynamics, namely: $\bar{\mathbf{n}}^{\text{stoc}} = \frac{1}{T} \sum_{t=0}^T \mathbf{n}(t)$. If the mean field description of the dynamics describes sufficiently well the underlying stochastic dynamics, by substituting the averaged configuration in equation (2) we should get $\mathbb{T}(\bar{\mathbf{n}}^{\text{stoc}}) \cdot \bar{\mathbf{n}}^{\text{stoc}} \simeq 0$. We find this to be the case for both models. We have therefore treated $\bar{\mathbf{n}}^{\text{stoc}}$ as our fixed points.

Through our procedure we want to study the stability in the neighbourhood of $\bar{\mathbf{n}}^{\text{stoc}}$, in order to predict the system's reaction to stochastic perturbations. To the extent that the mean field matrix correctly describes the system the metastable states will become unstable along directions in configuration space given by the eigenvectors \mathbf{e}_+ corresponding to eigenvalues with a positive real part $\text{Re}(\lambda) > 0$.

Once we know the form of the eigenspace we can monitor two important scalar quantities: the instantaneous distance from the fixed point

$$\delta n(t) = \|\delta \mathbf{n}(t)\| = \|\mathbf{n}(t) - \bar{\mathbf{n}}^{\text{stoc}}\| \quad (19)$$

and the maximum overlap between the perturbation and the eigenvectors $\{\mathbf{e}_+\}$ of the unstable subspace

$$Q(t) = \max \|\delta \mathbf{n}(t) \cdot \mathbf{e}_i\| \quad \forall \mathbf{e}_i \in \{\mathbf{e}_+\}. \quad (20)$$

The quantity in equation (19) tells us how far away the system is from the fixed point while the overlap in equation (20) tells us to what extent a deviation $\mathbf{n}(t) - \bar{\mathbf{n}}^{\text{stoc}}$ is within an unstable sub space. We expect $\delta n(t)$ to fluctuate around a low constant while $Q(t)$ is zero, since this would mean that the perturbations happen in the stable subspace, while a transition would induce a sudden increase in both $\delta n(t)$ and $Q(t)$.

Another way of picturing $Q(t)$ is as a measure of the activity of the occupancy on dangerous nodes. Indeed every non zero component of the unstable eigenvectors $\{\mathbf{e}_+\}$ will tell us which nodes of the interaction network are capable of pushing the system out of its metastable configuration. Namely if $e_+^j \neq 0$, where j indicates the component of the unstable eigenvector, this means the j th node is dangerous. The $Q(t)$ monitors the activity of such nodes. If one of these nodes were to become activated by mutations this would result in a rapid growth of $Q(t)$ and can be considered as a warning of a successive transition. In figure 3 of [15] it was discussed how these two quantities behave in the TaNa model and we demonstrated the forecasting power of the indicator $Q(t)$ and we gave an explanation on why we missed some of the transitions. Here we illustrate in figure 3 the temporal behaviour of $Q(t)$ and $\delta n(t)$ for both the TaNa model and the stochastic replicator system. The top panels contain weighted occupation plots while the bottom figures show the behaviour of the two quantities in $Q(t)$ and $\delta n(t)$. The arrow points at the new dangerous mutant that has entered the system, while the dashed vertical line indicates the moment it happens. Before the dashed line we can see how fluctuations in $\delta n(t)$ are bounded and $Q(t)$ is essentially equal to zero. After the dashed line, when the new mutant has entered the system, we see an explosion of both quantities.

We denote t^* the time at which the transition begins, which is set by the $\delta n(t)$ crossing a reasonably chosen threshold T_δ and staying consistently above this threshold (we have used $T_\delta = 150$ for the TaNa and $T_\delta = 0.05$ for the replicator model). Given the sharp increase of $\delta n(t)$ when approaching the transition, t^* does not depend strongly on the precise choice of

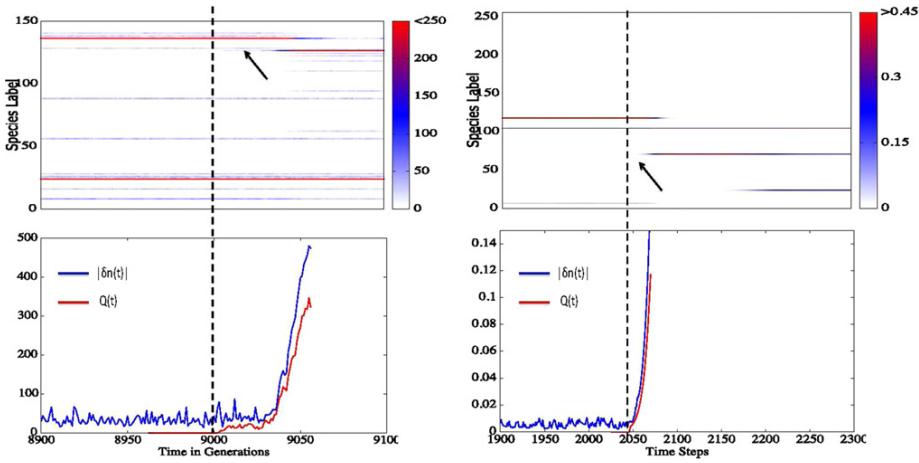


Figure 3. In the bottom panel of both figures we show the behaviour of $\delta n(t)$ (blue curve) and $Q(t)$ (red curve) while approaching the transition in the TaNa (left) and the replicator model with stochasticity (right). In the top panel, a weighted occupation plot is presented. We can see how the beginning of the transitions (dashed vertical black line) is triggered by a new mutant (black arrow) that quickly gains population. The arrival of the new dangerous mutant is singled by a peak in the $Q(t)$.

the threshold as long as its is chosen to be larger than the characteristic fluctuations of $\delta n(t)$ during the metastable configurations.

To define an alarm we determine an appropriate threshold A_Q on $Q(t)$. To do so we compare the number of false alarms with the number of missed transitions generated by different values of the chosen threshold A_Q . We define a false alarm when the $Q(t)$ crosses A_Q but then goes back under it before any transition occurs. On the other hand a missed transition corresponds to situations where $Q(t)$ remained below A_Q even though the given metastable configuration did become unstable and therefore a transition did occur.

In figure 4 we show these two quantities for different A_Q . The red curve is the fraction of missed transitions while the blue is the fraction of transitions that have produced false alarms. In the TaNa model, when increasing A_Q the fraction of false alarms decreases, as expected, while the fraction of missed transitions increases. The same figure for the replicator model shows how the procedure, although missing an increasing number of transitions, produce no false alarms at all.

The reason for this, we believe, has to do with the Langevin nature of the dynamics in the replicator model, i.e. deterministic dynamics + stochastic noise. Within this approach we expand the configuration vector $\mathbf{n}(t)$ in the \mathbb{M} 's eigenspace or generalised eigenspace plus noise. One obtains

$$\mathbf{n}(t) = \sum_k [c_k(0) \exp(\lambda_k t) \cdot \mathbf{e}_k + \epsilon_k] \quad (21)$$

where $c_k(0)$ are the coefficients of the expansion and ϵ_k is the noise. These dynamics are clearly dominated by those components for which $\text{Re}(\lambda_k) > 0$, but this is true only if $c_k(0) \neq 0$. When a node is populated by a mutation, in our framework this corresponds to setting $c_k(t) > 0$. From then on, the term is suppressed if and only if the ϵ_k points in the opposite direction at all times, which is highly unlikely. The same picture is less applicable to

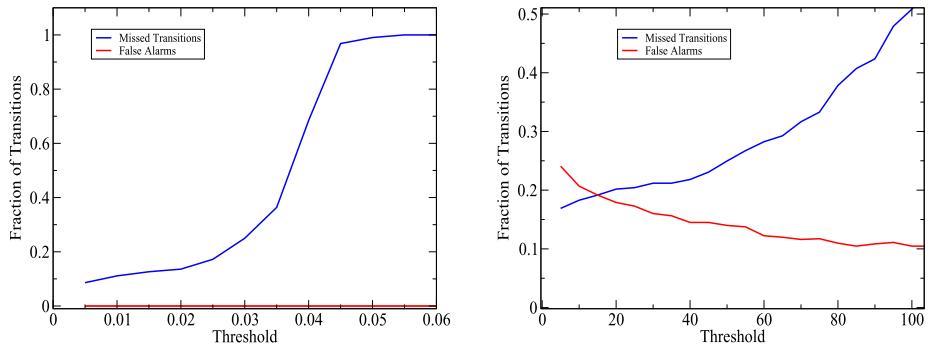


Figure 4. The behaviour of the fraction of false alarms and missed transitions for different values of alarm threshold A_Q in the replicator model (left panel) and the TaNa model (right panel). One can see how the procedure produced no false alarms in the replicator model which is consistent with what one might expect given the Langevin nature of the model.

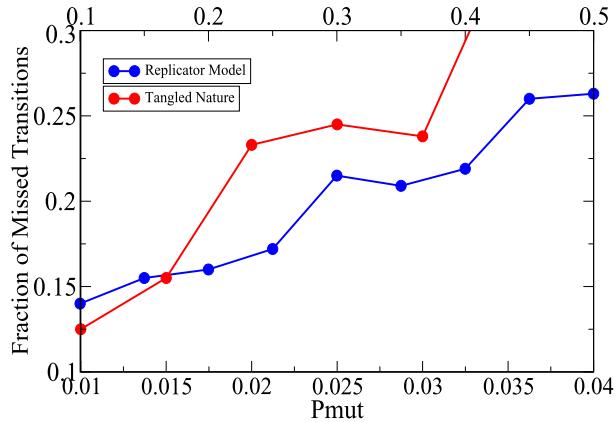


Figure 5. The fraction of missed transitions as a function of the noise in the system for both models. As the stochasticity is increased it becomes harder to forecast transitions in both models. The blue curve is the replicator model and it refers to the top x -axis, while the red curve is the TaNa model and refers to the bottom x -axis.

the TaNa model where all updates are stochastic and hence the separation into a robust deterministic part perturbed by a weak stochastic part is less applicable.

The way to interpret the missed transitions is to think of the fixed points as saddle points of a heterogeneous high dimensional energy landscape. The eigenspace of the mean field matrix tells us where the *downhill slopes* and *uphill barriers* are. Although it is far more likely for the system to leave the saddle point through a downhill slope, a stochastic perturbation may be able to push the system over a barrier. This interpretation is confirmed by figure 5 where we show that the fraction of missed transitions increases in both models as the degree of stochasticity is increased.

Once the threshold A_Q is fixed, we can determine the time t_{cross} at which $Q(t)$ goes above A_Q and determine the number of time steps $\Delta T = \|t^* - t_{\text{cross}}\|$ between the passing of the

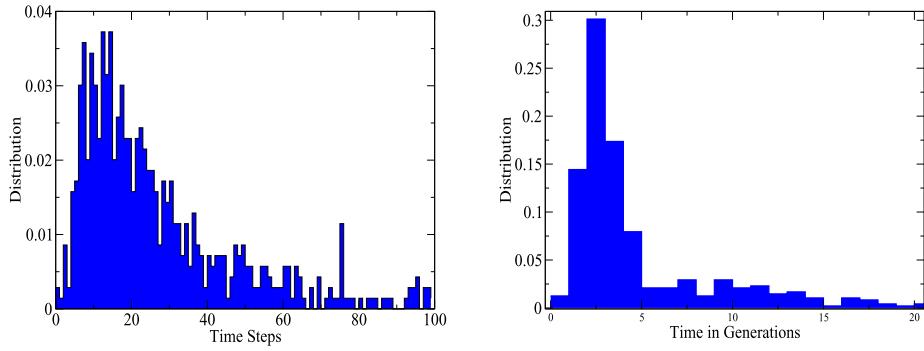


Figure 6. Distribution of the respite of the alarms for a given threshold. The left panel refers to the replicator model, for which $A_Q = 0.01$ and the right panel to the TaNa model, for which $A_Q = 20$.

threshold and the transition as given by the time t^* , at which the configuration starts to change significantly. In this way we can check the forecasting power of the indicator. In figure 6 we present the distribution of ΔT for $A_Q = 0.01$ and $A_Q = 20$ respectively for the replicator and the TaNa model. We can see that in the replicator model the crossing times are tenths of time steps before the transition time. This means that the system will go through many cycles of updates before the transition occurs. In the TaNa model in more than 50% of cases $\Delta T \in [2, 5]$. As explained above when introducing the model, one generation corresponds to the average number of time steps necessary to remove everyone from the system, i.e. $\frac{N(t)}{p_{\text{kill}}}$ individual updates. So even low values of ΔT will involve many individual updates and hence can be considered to correspond to a strong forecasting power.

5. Incomplete knowledge

An obvious shortcoming concerning the real-life application of the forecasting procedure described so far is that we make use of complete knowledge of the entire space of agents and their interactions, i.e. we use both the actually realised and the potential part of the space of agents. In this section we first consider how the lack of full knowledge of the interaction strength between agents influences our ability to detect approaching transitions. We next consider a much simpler measure than the overlap function $Q(t)$. This new measure is inspired by the analysis presented above and leading to $Q(t)$ but avoids access to information about the adjacent possible, i.e. information about agents that are not extant in the system at the time of forecasting. Our new measure only makes use of the time evolution of directly observable quantities.

5.1. Error in interactions

We investigate the effect of lack of complete information concerning the iterations between agents by introducing an error in the interaction matrix used for the mean field treatment. We do this in the following way

$$J_{ij}^e = J_{ij}^{\text{sim}} + \chi \quad (22)$$

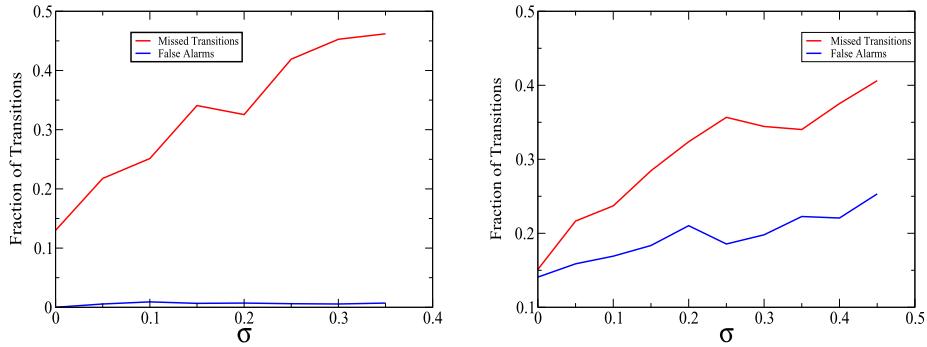


Figure 7. The fraction of the missed transitions and the fractions of false positive as function of the σ of the distribution of the random error in the interactions. Once again we have used $A_Q = 30$ for the TaNa (right panel) and $A_Q = 0.01$ for the replicator model (left panel).

where χ is $N(0, \sigma)$, i.e. a normally distributed random variable, of mean 0 and variance σ . We then repeat the exact same procedure outlined in the previous section but using \mathbf{J}^e in the forecasting calculations while the simulations evolves according to \mathbf{J}^{sim} .

In figure 7 we present the fractions of transitions we are not able to forecast (missed transition) and the fractions of false alarms we generate as function of the variance σ , i.e. as a function of how much the interaction matrix used for the stability analysis differs from the correct set of interactions. For the TaNa model (see the right panel) we can notice that for $\sigma < 0.2$ we are still able to forecast around 70% of the transitions and we generate less than 20% of false alarms. This is an encouraging result since a $\sigma = 0.2$ is clearly a significant error given that $J_{ij} \in (-1, 1)$. A very similar result holds for the replicator model.

5.2. New procedure

We now discuss a forecasting procedure that does not need any knowledge about “in potentia” agents. We only need to focus on the highly occupied nodes present in the system. We only know what we see without making any use of the non active part of the interaction network, nor of the poorly occupied nodes. By applying the LSA to the occupied network we can check that, during a stable phase, the configuration corresponds to a situation where the spectrum of the stability matrix \mathbb{M} consists of eigenvalues that all have negative real parts. As the system evolves new mutants appear. As an indicator of approaching transitions we track the growths of the occupancy of these new agents, if their occupancy exceeds a certain threshold we check the spectrum of the updated \mathbb{M} , in which the new agents are included. In case the spectrum now includes a positive eigenvalue we take this as an indicator of an approaching transition out of the present metastable configuration. This will be our new alarm. That rapidly growing new types may destabilise the configurations of the TaNa model was also discussed by Becker and Sibani [36].

In figure 8 we show the results of an application of this new procedure. In both panels the red vertical lines indicate the times of appearance of a species able to change the stability of the system. We can qualitatively see from the figure that just after the alarms the system actually undergoes a transition. In the left panel of figure 8 the blue curves represent the frequencies of the most occupied strategies in the replicator model. We can see how directly after the red lines, the alarm times, a new strategy starts gaining frequency and eventually puts

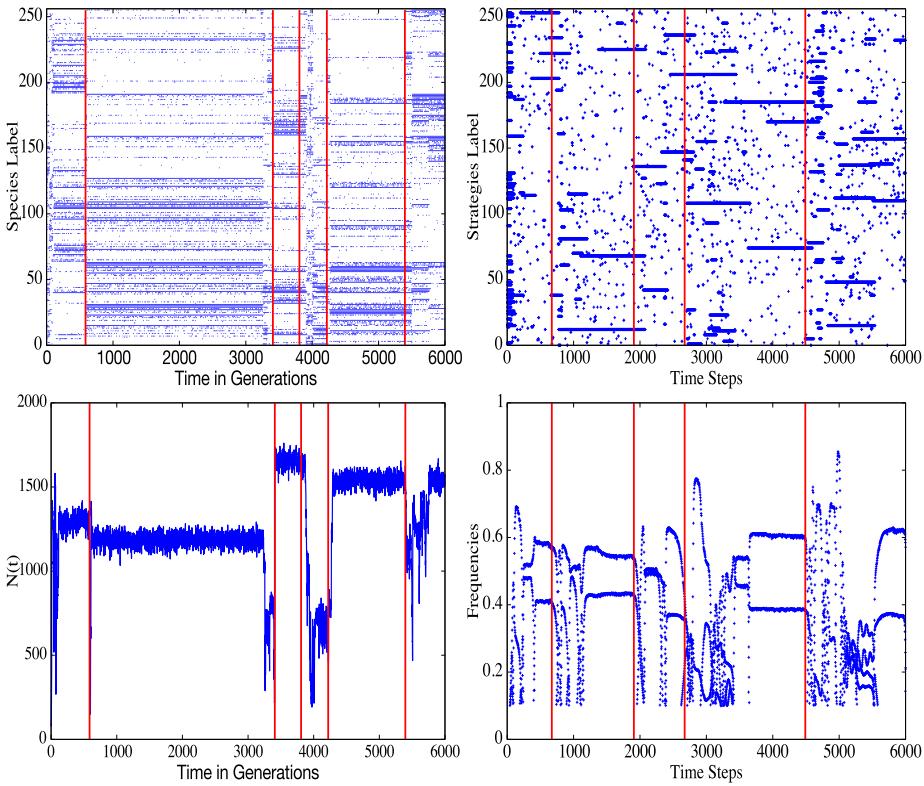


Figure 8. Left panel: occupancy distribution of the types. The genotypes are labelled arbitrarily and a dot indicates a type which is occupied at the time t . The punctuated dynamics is clearly visible: quasi-stable periods alternate with brief periods of hectic transitions. Right panel: the frequencies of the strategies. Each colour belongs to a different strategy. Once again the transitions from one meta stable configuration (approximate fixed point) to another is clear.

an end to the stable configuration. It should be stressed that using the growth in the population of the new mutant alone as an early warning signal will not work, since it would produce a very high and continuous rate of false alarms. It is the combination of the spectral analysis with the monitoring of the arrival of new mutants that makes the method very powerful.

In the bottom left panel we show the total number of individuals present in the TaNa model: $N(t) = \sum_j n_j(t)$. A transition to a new metastable configuration is associated with a sudden change of this quantity. We notice that after each alarm $N(t)$ exhibit a significant change induced by the arrival of the new fast growing type. Preliminary analysis indicates that this procedure is able to forecast transitions with an accuracy similar the $Q(t)$ indicator. Further investigation of the efficiency and reliability of using the growth of newcomers as indicators of approaching transitions is underway. Obviously this can make our procedure more readily applicable to real systems since we would then only need directly observable information.

6. Summary and conclusion

We have described a new procedure for forecasting transitions in high dimensional systems with stochastic dynamics. Our method is of relevance to systems where the macroscopic dynamics at the systemic level is not adequately captured by a well defined set of essentially deterministic collective variables (e.g. as handled by Langevin equations). Hence we are dealing with situations which are not captured by the application of bifurcation theory such as those considered by Scheffer and collaborators [2, 3, 6]. We have in mind complex systems in which the dynamics involve some evolutionary aspects, in particular situations where the dynamics generates new degrees of freedom. Examples of these include biological evolution, or economical and financial systems, where new agents (organisms, strategies or companies, say) are produced as an intrinsic part of the dynamics. We have demonstrated by use of two models of varying degree of stochasticity (the Tangled Nature model and the stochastic replicator model) where a combination of analytic linear stability analysis and simulation allows one to construct a signal (overlap with unstable directions) which can be used to forecast a very high percentage of all transitions.

The weakness of the procedure is that for real situations of interest (e.g. an ecosystem or a financial market) one may obviously not possess complete information. One will typically not have access to all the information about the interaction amongst the agents. This turns out to be less of a problem, since we can show that even with a 10% inaccuracy in interaction strengths, we are still able to forecast a substantial percentage of transitions. Another shortcoming is that in real situations it can also be very difficult to know the nature of the new agents that may arrive as the system evolves. Our full mathematical procedure suggests a way to overcome this problem. Namely, the eigenvector analysis showed that transitions are often accompanied by the arrival of new agents, which exhibit a rapid growth in their relative systemic weight. We found that simply monitoring the rapidly growing new agents can enable prediction of major systemic upheavals. I.e. approaching transitions might not be apparent by focusing on the systemic heavyweights, but rather one should keep a keen eye on the tiny components to monitor whether they suddenly start to flourish. This can often be the signal of upcoming systemic changes.

To address on a real system how crucial incomplete data and limited model accuracy is, our next step is to test our approach on real data streams including a high frequency financial time series.

Acknowledgments

This work was supported by the European project CONGAS (Grant FP7-ICT-2011-8-317672). We are grateful for computer time on Imperial College's High Performance Cluster. JG is a Postdoctoral Fellow of the Research Foundation—Flanders.

References

- [1] Sibani P and Jensen H J 2013 *Stochastic Dynamics of Complex Systems* (London: Imperial College Press)
- [2] Scheffer M, Carpenter S R, Lenton T M, Bascompte J, Brock W, Dakos V, van de Koppel J, van de Leemput I A, Levin S A, van Nes E H *et al* 2012 Anticipating critical transitions *Science* **338** 344–8
- [3] Scheffer M, Bascompte J, Brock W A, Brovkin V, Carpenter S R, Dakos V, Held H, Van Nes E H, Rietkerk M and Sugihara G 2009 Early-warning signals for critical transitions *Nature* **461** 53–9

- [4] Kuehn C 2011 A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics *Physica D: Nonlinear Phenomena* **240** 1020–35
- [5] Kuehn C 2013 A mathematical framework for critical transitions: Normal forms, variance and applications *Journal of Nonlinear Science* **23** 457–510
- [6] Scheffer M 2009 *Critical Transitions in Nature and Society* (Princeton, NJ: Princeton University Press)
- [7] Lenton T M 2011 Early warning of climate tipping points *Nat. Clim. Change* **1** 201–9
- [8] Lenton T M, Held H, Kriegler E, Hall J W, Lucht W, Rahmstorf S and Schellnhuber H J 2008 Tipping elements in the Earth's climate system *Proc. Natl Acad. Sci.* **105** 1786–93
- [9] Veraart A J, Faassen E J, Dakos V, van Nes E H, Lürling M and Scheffer M 2012 Recovery rates reflect distance to a tipping point in a living system *Nature* **481** 357–9
- [10] Carpenter S R and Brock W A 2006 Rising variance: a leading indicator of ecological transition *Ecol. Lett.* **9** 311–8
- [11] Battiston S, Farmer J D, Flache A, Garlaschelli D, Haldane A G, Heesterbeek H, Hommes C, Jaeger C, May R and Scheffer M 2016 Complexity theory and financial regulation *Science* **351** 818–9
- [12] Wang R, Dearing J A, Langdon P G, Zhang E, Yang X, Dakos V and Scheffer M 2012 Flickering gives early warning signals of a critical transition to a eutrophic lake state *Nature* **492** 419–22
- [13] Dakos V, van Nes E H and Scheffer M 2013 Flickering as an early warning signal *Theor. Ecol.* **6** 309–17
- [14] Kuehn C, Zschaler G and Gross T 2014 Early warning signs for saddle-escape transitions in complex networks arXiv:1401.7125
- [15] Caroli A, Piovani D and Jensen H J 2014 Forecasting transitions in systems with high dimensional stochastic complex dynamics: A linear stability analysis of the tangled nature model *Phys. Rev. Lett.* **113** 264102
- [16] Christensen K, di Collobiano S A, Hall M and Jensen H J 2002 Tangled nature model: A model of evolutionary ecology *J. Theor. Biol.* **216** 73–84
- [17] Hall M, Christensen K, di Collobiano S A and Jensen H J 2002 Time-dependent extinction rate and species abundance in a tangled-nature model of biological evolution *Phys. Rev. E* **66** 011904
- [18] Anderson E P and Jensen H J 2005 Network properties, species abundance and evolution in a model of evolutionary ecology *J. Theor. Biol.* **232** 551–8
- [19] Lawson D and Jensen H J 2006 The species-area relationship and evolution *J. Theor. Biol.* **241** 590–600
- [20] Sigurd D and Opper M 1989 Replicators with random interactions: A solvable model *Phys. Rev. A* **39** 4333
- [21] Tokita K and Yasutomi A 2003 Emergence of a complex and stable network in a model ecosystem with extinction and mutation *Theor. Popul. Biol.* **63** 131–46
- [22] Hadeler K P 1981 Stable polymorphisms in a selection model with mutation *SIAM J. Appl. Math.* **41** 1–7
- [23] Hofbauer J and Sigmund K 1998 *Evolutionary Games and Population Dynamics* (Cambridge: Cambridge University Press)
- [24] Nowak M A, Komarova N L and Niyogi P 2001 Evolution of universal grammar *Science* **291** 114–8
- [25] Eigen M 1971 Self-organization of matter and the evolution of biological macromolecules *Naturwissenschaften* **58** 465
- [26] Eigen M and Schuster P 1977 A principle of natural self-organization *Naturwissenschaften* **64** 541–65
- [27] Lauring A S and Andino R 2010 Quasispecies theory and the behavior of rna viruses *PLoS Pathog.* **6** e1001005
- [28] Taylor P D and Jonker L B 1978 Evolutionary stable strategies and game dynamics *Math. Biosci.* **40** 145–56
- [29] Smith J M 1982 *Evolution and the Theory of Games* (Cambridge: Cambridge University Press)
- [30] Lotka A J 1910 Contribution to the theory of periodic reactions *J. Phys. Chem.* **14** 271–4
- [31] Goodwin R M 1982 A growth cycle *Essays in Economic Dynamics* (New York: Springer) 165–70
- [32] Hofbauer J 1981 On the occurrence of limit cycles in the volterra-lotka equation *Nonlinear Anal. Theory Methods Appl.* **5** 1003–7
- [33] Price G R 1970 Selection and covariance *Nature* **227** 520–1

- [34] Hamilton W D 1970 Selfish and spiteful behaviour in an evolutionary model *Nature* **228** 1218–20
- [35] Page K M and Nowak M A 2002 Unifying evolutionary dynamics *J. Theor. Biol.* **219** 93–8
- [36] Becker N and Sibani P 2004 Evolution and non-equilibrium physics: A study of the tangled nature model *Europhys. Lett.* **105** 18005–11

Bibliography

- [1] V. Ambegaokar, B.I. Halperin, D.R. Nelson, and E.D. Siggia. Dissipation in two-dimensional superfluids. *Phys. Rev. Lett.*, 40:783–786, 1978.
- [2] V. Ambegaokar, B.I. Halperin, D.R. Nelson, and E.D. Siggia. Dynamics of superfluid films. *Phys. Rev. B*, 21:1806–1826, 1980.
- [3] P.W. Anderson. More is different. *Science*, 177:393–396, 1972.
- [4] P.W. Anderson. *Basiv Notions of Condensed Matter Physics*. The Benjamin/Cummings Publishing Company, Inc., Menlo Park, California, 1984.
- [5] A.-L. Barabási and H.E. Stanley. *Fractal Concepts in Surface Growth*. Cambridge University Press, Cambridge, 1995.
- [6] M.R. Beasley, J.E. Mooij, and T.P. Orlando. Possibility of vortex-antivortex pair dissociation in two-dimensional superconductors. *Phys. Rev. Lett.*, 42:1165–1168, 1979.
- [7] N. Becker and P. Sibani. Evolution and non-equilibrium physics: A study of the tangled nature model. *EPL*, 105:18005, 2014.
- [8] J.J. Binney, N.J. Dowrick, A.J. Fisher, and M.E.J. Newman. *The Theory of Critical Phenomena. An Introduction to the Renormalization Group*. Oxford University Press, Oxford, 1992.
- [9] S. Boettcher and A. Percus. Optimization with extremal dynamics. *Phys. Rev. Lett.*, 86:5211–5214, 2001.
- [10] S. Camazine, J.-L. Deneubourg, N.R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau. *Self-organization in Biological Systems*. Princeton University Press, 2001.

- [11] V. Cataudella and P. Minnhagen. Simple estimate for vortex fluctuations in connection with high- t_c superconductors. *Physica C*, 166:442–450, 1990.
- [12] P.M. Chaikin and T.C. Lubensky. *Principles of condensed matter physics*. Cambridge University Press, Cambridge, 1995.
- [13] K. Christense and N.R Moloney. *Complexity and Criticality*. Imperial College Press, 2005.
- [14] K. Christensen, S. A. di Collobiano, M. Hall, and H. J. Jensen. Tangled nature model: A model of evolutionary ecology. *J. Theor. Biol.*, 216:73–84, 2002.
- [15] S. Clar, B. Drossel, and F. Schwabl. Scaling laws and simulation results for the self-organized critical forest-fire model. *Phys. Rev. E*, 50:1009, 1994.
- [16] K. Cristensen, H. Flyvbjerg, and Z. Olami. Self-organized critical forest-fire model: Mean-field theory and simulation results in 1 to 6 dimensions. *Phys. Rev. Lett.*, 71:2737, 1993.
- [17] K Cristensen and Z. Olami.
- [18] S. Doniach and B.A. Huberman. Topological excitations in two-dimensional superconductors. *Phys. Rev. Lett.*, 42:1169–1172, 1979.
- [19] B. Drossel and F. Schwabl. *Phys. Rev. Lett*, page 1629.
- [20] J. Feder. *Fractals*. PlenumPress., 1988.
- [21] William Feller. *An Introduction to Probability Theory and its Applications, vol. II*. John Wiley, New York London Sidney Toronto, 1966.
- [22] R.A. Fisher and L.H.C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Camb. Phil. Soc.*, 24:180–190, 1928.
- [23] M. Fréchet. Sur la loi de probabilit de lcart maximum. *Ann. Soc. Polon. Math.*, 6:93, 1927.
- [24] J Galambos. *The Asymptotic Theory of Extreme Order Statistics*. John Wiley & Sons, New York, 1978.
- [25] David E. Goldberg. *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Reading, Massachusset, 1989.

- [26] P. Grassberger. *J. Phys. A*, page 2081.
- [27] G. Grinstein, T. Hwa, and H.J. Jensen. $1/f^\alpha$ noise in dissipative transport. *Phys. Rev. A*, 45:R559–R562, 1992.
- [28] E J Gumbel. *Statistics of Extremes*. Columbia University Press, 1958.
- [29] Kantz H. and Schreiber T. *Nonlinear Time Series Analysis*. Cambridge University Press, 2000.
- [30] M. Hall, K. Christensen, S. A. di Collobiano, and H. J. Jensen. Time-dependent extinction rate and species abundance in a tangled-nature model of biological evolution. *Phys. Rev. E*, 66, 2002.
- [31] Theodore E Harris. *The Theory of Branching Processes*. Dover, 1989.
- [32] P.G. Higgs and B. Derrida. Genetic distance and species formation in evolving populations. *J. Mol. Evolution*, 35(5):454–465, 1992.
- [33] B. Hu and J. Rudnick. Exact solutions to the feigenbaum renormalization-group equations for intermittency. *Phys. Rev. Lett.*, 48:1645, 1982.
- [34] B.A. huberman, R.J. Myerson, and S. Doniach. Dissipation near the critical point of a two-dimensional superfluid. *Phys. Rev. Lett.*, 40, 1978.
- [35] H.J. Jensen. *Self-Organized Criticality. Emergent Complex Behavior in Physical and Biological Systems*. Cambridge University Press, 1998.
- [36] H.J. Jensen and H. C. Fogedby. Phonon-kink interference in the ϕ^4 model. *Phys. Scr.*, 31:210–214, 1985.
- [37] H.J. Jensen and H. Weber. A phenomenological study of vortices in a two dimensional xy-model in a magnetic field. *Phys. Rev. B.*, 45:10468–10472, 1992.
- [38] L.P. Kadanoff. *Statistical Physics. Statics, Dynamics and Renormalization*. World Scientific Publishing Co., Singapore, 2000.
- [39] A.M. Kadin, K. Epstein, and A.M. Gildman. Renormalization and the kosterlitz-thouless transition in a two-dimensional superconductor. *Phys. Rev. B*, 27:6691–6702, 1983.
- [40] J. M. Kosterlitz and D. J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys. C: Solid State Phys.*, 6:1181–1203, 1973.

- [41] J.M. Kosterlitz. The critical properties of the two-dimensional xy model. *J. Phys. C: Solid State Phys.*, 7:1046–1060, 1974.
- [42] S. Kotz and S. Nadarajah. *Extreme Value Distributions: Theory and Applications*. Imperial College Press, 2000.
- [43] Joachim Krug. Records in a changing world. *Journal of Statistical Mechanics, Theory and Experiment*, page P07001, 2007.
- [44] Shang-Keng Ma. *Statistical Mechanics*. World Scientific, Singapore, 1985.
- [45] E. Milotti. 1/f noise: a pedagogical review. <http://arxiv.org/pdf/physics/0204033>.
- [46] P. Minnhagen. The two-dimensional coulomb gas, vortex unbinding, and superfluid-superconducting films. *Rev. Mod. Phys.*, 59:1001–1066, 1987.
- [47] P. Minnhagen and P. Olsson. Monte carlo calculation of the vortex interaction for high- t_c superconductors. *Phys. Rev. B*, 44:4503–4511, 1991.
- [48] N. Glick. Breaking records and breaking boards. *The American mathematical monthly*, 85:2–26, 1978.
- [49] David R. Nelson. *Defects and Geometry in Condensed Matter Physics*. Cambridge University Press, Cambridge, 2002.
- [50] V. B. Nevezorov. *Records: Mathematical Theory*. American Mathematical Society, 2001.
- [51] A.E Nicholson and P. Sibani. Cultural evolution as a nonstationary stochastic process. *Complexity*, page DOI: 10.1002/cplx.21681, 2015.
- [52] P. Olsson and P. Minnhagen. On the helicity modulus, the critical temperature and monte carlo simulations for the two-dimensional xy-model. *Phys. Scr.*, 43:203–209, 1991.
- [53] Per Bak, Chao Tang and Kurt Wiesenfeld. Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters*, 59:381–384, 1987.
- [54] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Commun. Math. Phys.*, 74:189–197, 1980.
- [55] W.H. Press. Flicker noises in astronomy and elsewhere. *Comments on Modern Physics, Part C - Comments on Astrophysics*, 7:103–119, 1978.

- [56] I. Procaccia and H.G.Schuster. Functional renormalization-group theory of universal 1/f noise in dynamical systems. *Phys. Rev. A*, 28:1210, 1983.
- [57] G. Pruessner. *Self-Organised Criticality. Theory, Models and Characterisation*. Cambridge University Press, 2011.
- [58] G. Pruessner and H.J. Jensen. Broken scaling in the forest-fire model. *Phys. Rev. E*, 65:056707, 2002.
- [59] F. Reif. *Fundamentals of Statistical and Thermal Physics*. McGraw-Hil, New York, 1965.
- [60] C. Rhodes and R. Anderson. *Nature*, page 600.
- [61] Per Arne Rikvold and R. K. P. Zia. Punctuated equilibria and 1/f noise in a biological coevolution model with individual-based dynamics. *Phys. Rev. E*, 68, 2003.
- [62] R.Solé and B. Goodwin. *Signs of Life, How Complexity Pervades Biology*. Basic Books, New York, 2000.
- [63] Walter Rudin. *Real and complex analysis*. Mc Graw Hill, New York, 1966.
- [64] T.C. Schelling. Dynamic models of segregation. *J. Math. Sociol.*, 1:143–186, 1971.
- [65] H.G. Schuster and W. Just. *Deterministic Chaos. An Introduction*. Wiley-VCH Publishers, 2005.
- [66] P. Sibani, M. Brandt, and P. Alstrøm. Evolution and extinction dynamics in rugged fitness landscapes. *Int. J. Modern Phys. B*, 12:361–391, 1998.
- [67] P. Sibani and H.J. Jensen. *Stochastic Dynamics of Complex Systems*. Imperial College Press, 2013.
- [68] P. Sibani and Littlewood P.B. Slow dynamics from noise adaptation. *Phys. Rev. Lett.*, 71:1482, 1993.
- [69] J. Maynard Smith. *Evolution and the Theory of games*. Cambridge University Press, 1982.
- [70] D. Sornette. *Critical Phenomena in Natural Sciences. Chaos, Fractals, Selforganization and Disorder: Concepts and Tools*. Springer-Verlag, Berlin, 2004.
- [71] Dietrich Stauffer and Amnon Aharony. *Introduction to Percolation Theory*. Taylor & Francis, 1994.

- [72] E. Tagliazucchi, P., Pablo Balenzuela, D. Daniel Fraiman, and D.R. Chialvo. Criticality in large-scale brain fmri dynamics unveiled by a novel point process analysis. *frontiers in Physiology*, 3:15, 2012.
- [73] M. Tinkham. *Introduction to Superconductivity*. Dover Publications, 2004. Second revised edition.
- [74] F. Varela, J-P Lachaux, E Rodriguez, and J Martierie. The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2:229–239, 2001.
- [75] P. Vazquez, J.A. del Rio, K.G. Cedano, M. Martinez, and H. J. Jensen. An entangled model for sustainability indicators. *PLoS ONE*, 10:e0135250, 2015. doi:10.1371/journal.pone.0135250.
- [76] H. Weber and H.J. Jensen. Crossover from three- to two-dimensional behavior of the vortex energies in layered xy models for high- t_c superconductors. *Phys. Rev. B*, 44:454–457, 1991.
- [77] M.B. Weissman. $1/f$ noise and other slow, nonexponential kinetics in condensed matter. *Rev. Mod. Phys.*, 60:537–571, 1988.
- [78] K. Wilson. The renormalisation group and critical phenomena. http://nobelprize.org/nobel_prizes/physics/laureats/1982/wilson-lecture.html, 1982.
- [79] R K P Zia and Per Arne Rikvold. Fluctuations and correlations in an individual-based model of evolution. *J. Phys. A*, 37:5135–5155, 2004.