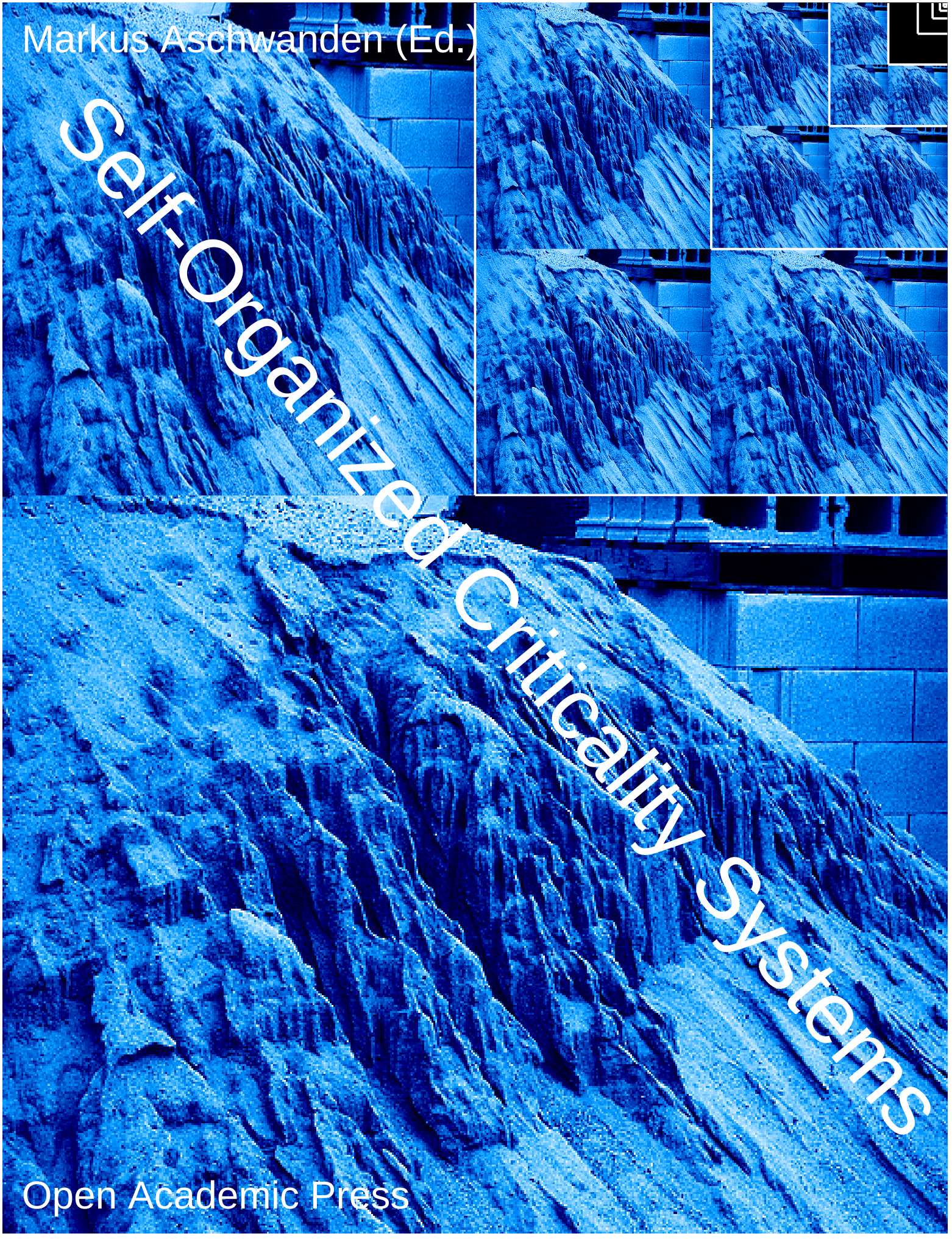

Markus Aschwanden (Ed.)
Self-Organized Criticality Systems
Open Academic Press

Markus J. Aschwanden (Editor)

Open Academic Press Berlin Warsaw 2013

Self-Organized Criticality Systems

Dr. Markus Josef Aschwanden (Editor)
Lockheed Martin, Advanced Technology Center
Solar & Astrophysics Lab, Dept., Org. ADBS, Bldg.252
3251 Hanover Street, Palo Alto, CA 94304, USA
Telephone: 650-424-4001, Fax: 650-424-3994
E-mail: aschwanden@lmsal.com

Published by Open Academic Press Berlin Warsaw 2013
Open Academic Press, Großbeerenstraße 2-10, 12107 Berlin, Germany
Open Academic Press/Villa Europa, ul. Sienna 64, 00-807 Warsaw, Poland

Contents

Chapter 1
Introduction

Norma B. Crosby

> In turn, these invariance properties suggest that the presence of scaling distributions in data obtained from complex natural or engineered systems should be considered the norm rather than the exception and should not require "special" explanations.
>
> Willinger et al. (2004)

1.1 A New Theory Emerges

In the mid-1980s a new theory emerged that aimed to explain how complex nonlinear systems with many degrees of freedom observed in nature are able to produce powerlaw relationships from simple redistribution rules of nearest-neighbor interactions. The incentive behind this theory was that it could be the underlying concept for temporal and spatial scaling observed in a wide class of dissipative systems with extended degrees of freedom. It was the Danish theoretical physicist Per Bak and his co-authors who introduced this concept that became known as SOC (Bak et al. 1987; 1988). This acronym stands for Self-Organized Criticality and since its birth this concept has been applied to a wide range of disciplines covering solar physics, astrophysics, magnetospheric physics, geophysics, biophysics, and social sciences.

Many systems consist of a large number of entities that interact in a complex way and exhibit nonlinear behavior; they are called nonlinear dissipative systems. Indeed, the Solar System is full of multi-scale phenomena that obey nonlinear spatio-temporal scaling laws. On Earth, such extreme nonlinear events are known as earthquakes, landslides, wildfires, volcanoes, snow avalanches, rock-falls, crashes in the stock market, etc. Their counterparts

Norma B. Crosby
Belgian Institute for Space Aeronomy, Ringlaan-3-Avenue Circulaire, B-1180 Brussels, Belgium, e-mail: Norma.Crosby@aeronomie.be

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

in space - solar storms - range from solar flares, coronal mass ejections, substorms in the magnetosphere to solar energetic particle events. What do all these natural dynamic phenomena (examples are displayed in Fig. 1.1), have in common?

1. They cover a large range of temporal, as well as spatial scales.
2. The most extreme events, known as "black swans" (Taleb 2007), are of concern to society.
3. There are large databases so that statistical approaches can be used for interpreting the data characterizing the phenomena.
4. Size distributions (on log-scale) of parameters describing the phenomena (volumes, energies, etc.) cover many orders of magnitude.
5. Powerlaw-like behavior has been found to be a universal characteristic of such phenomena.

Though it is most often the largest "avalanche" events that make the headlines in the newspapers, the myriads of smaller events share the same statistical properties. With databases becoming larger and larger (covering long time spans) it is not always possible to study each single event and a statistical approach such as a frequency distribution can be used instead. Powerlaw behavior is systematically observed when frequency distributions are constructed of both measured (e.g. peak count rate, total duration) and theoretical (e.g. total energy released) parameters describing the events (identified as "avalanches" in this Chapter) constituting the "avalanche" database. Performing a frequency distribution on an "avalanche" database is a valid approach if the phenomenon being studied results from the same mechanisms of energy release on all scales.

Why does nature produce powerlaw behavior? What is (are) the mechanism(s) responsible for powerlaws observed in nature as well as in social sciences? Various concepts (models) have been proposed to explain this observed powerlaw signature. One of these is SOC that characterizes the behavior of dissipative systems that contain a large number of elements interacting via nearest-neighbor interactions over short and long ranges. The systems evolve to a critical state in which a minor event starts a chain reaction that can affect any number of elements in the system. Frequency distributions of the output parameters from the chain reaction taken over a period of time can be represented by powerlaws.

Several reviews, textbooks, and monographs have been written on SOC regarding phenomena that display this behavior as well as models that simulate SOC; see for example Bak (1996), Jensen (1998), Turcotte (1999), Charbonneau et al. (2001), Hergarten (2002), Sornette (2004), Christensen and Moloney (2005), Aschwanden (2011), Crosby (2011), and Pruessner (2012). In the current book an inter-disciplinary approach is applied to the SOC concept. The book covers all types of naturally occurring "avalanche" phenomena where SOC behavior has been studied and presents both observational results as well as results obtained by theoretical models.

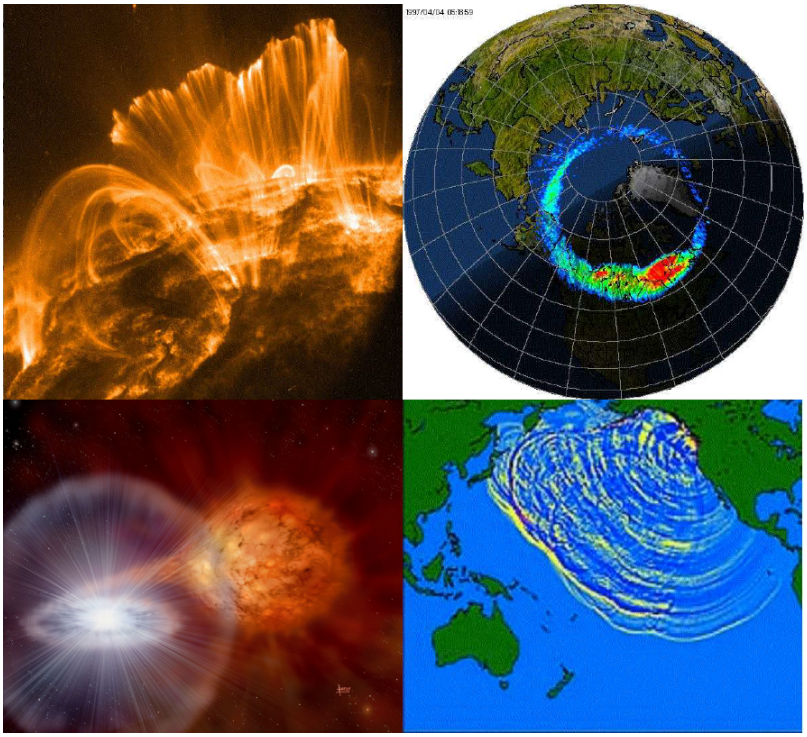


Fig. 1.1: Examples of natural phenomena in space and on Earth that exhibit powerlaw behavior. Upper left-hand corner: Solar flare of 2000 Nov. 9 observed in EUV with the TRACE spacecraft in 171 $\AA$ (credit: NASA, TRACE), upper right-hand corner: Global image of the auroral oval observed by the Ultraviolet Imager (UVI) onboard the NASA satellite "Polar" (credit: NASA, Polar/UVI Team, George Parks), lower left-hand corner: Artistic rendering of the cataclysmic variable star RS Ophiuchi, which exhibits a nova outburst about every 20 years. This binary system contains a white dwarf and a red giant with mass transfer (credit: PPARC, David A. Hardy), lower right-hand corner: Satellite recording of tsunami waves produced by one of the 10 largest earthquakes, originating in North America (credit: NOAA).

In Section 2 of this Chapter it is described how frequency distributions are applied as a statistical tool. This is followed by an introduction to the SOC concept and to the models that have been built to simulate powerlaw behavior. Examples of SOC observed in natural phenomena both in space and on Earth are presented in Section 4. The end of the Chapter asks the question: What does it all mean?

1.2 Frequency distribution: a powerful tool

A frequency distribution, also known as log Number vs. log Size (log N - log S) diagram, size distribution, or occurrence frequency distribution, is a function that describes the occurrence rate of events as a function of their size. It is usually plotted as a histogram of the logarithmic number versus the logarithmic size. Input for a frequency distribution is a database (list or catalogue) of "avalanches" characterized by some size parameter. There are essentially two ways to construct a log-log histogram from a database:

1. Logarithmically binned histogram if large statistics is available (n $\geq 10^2$, ..., 10^3).
2. Rank-order plot if the size of the statistical sample is rather small.

For numerous natural phenomena it is found that the differential frequency distributions of the output parameters describing a given phenomenon taken over a period of time can be represented by powerlaws of the form:

$$N(x)dx = N_0 x^{-\alpha_x} dx \tag{1.1}$$

where N(x) is the number of events recorded with the parameter x of interest in a "differential" bin, and N_0 is a constant.

For the cumulative frequency distribution, it is the integral which expresses in each bin the sum of all events that are larger than the size parameter of the bin x:

$$N^{cum}(> x) = \int_x^\infty x^{-\alpha_x} dx \propto x^{-(\alpha_x - 1)} \tag{1.2}$$

If the differential frequency is a powerlaw function with slope α, the cumulative frequency distribution is expected to have a flatter powerlaw slope by one (α-1).

Fig. 3.1 is an example of a differential frequency distribution performed on the peak count rate solar flare data recorded by the Wide Angle Telescope for Cosmic Hard X-Rays (WATCH) experiment aboard the GRANAT satellite. WATCH measures photons in the deka-keV range and the WATCH database covers 2.5 years of observation (1990 to mid-1992). Fig. 1.3 shows the time profiles of three WATCH flares observed between 15:00 and 16:20 UT on 1990 June 19. The three peak count rates that are identified as "a", "b" and "c" in Fig. 1.3 are included in the Fig. 3.1 statistics. The distribution displayed in Fig. 3.1 follows a powerlaw with a slope of -1.58 $\pm$ 0.02 extending for almost three orders of magnitude (Crosby et al. 1998). The turn-over in the lower end of the frequency distribution may be attributed to detector sensitivity (missing the small events in the background noise).

For some "avalanche" parameters exponential turn-overs in the upper end of the frequency distribution are observed and may be due to two important issues:

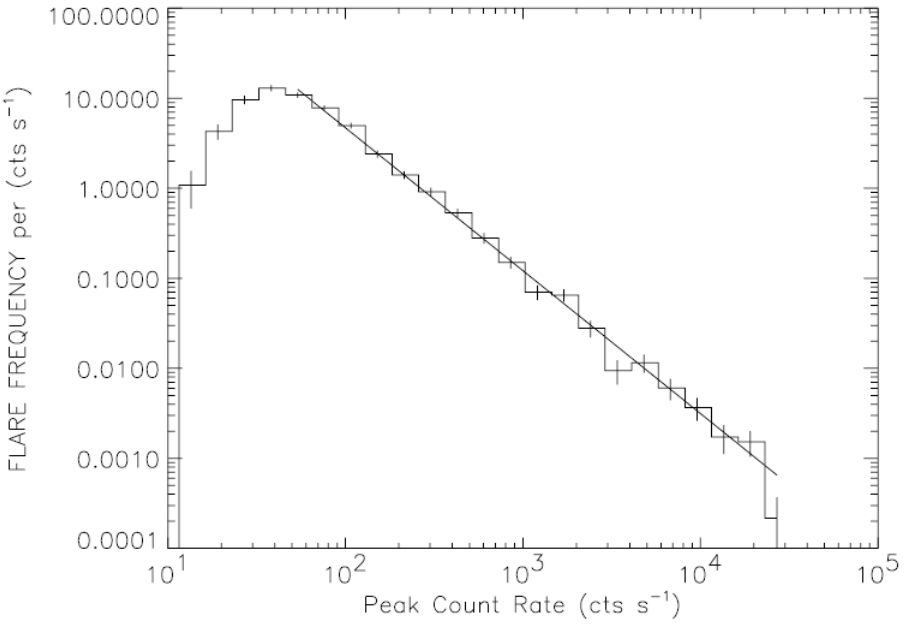


Fig. 1.2: The frequency distribution on the WATCH solar flare peak count rate data for the total observing period (1990 - mid-1992). It is well-represented by a powerlaw with a slope of -1.58 $\pm$ 0.02 extending over almost three orders of magnitude (Crosby et al. 1998).

1. Length of the dataset (missing of long-term statistics): observations have not been performed over a long enough period of time to cover all the statistics of the "avalanches".
2. Physical limit to the size of a "black swan" event: "Finite-size effects" ("avalanches" reach the boundaries of a system; Chapman et al. (1998) discusses this effect in regard to substorms).

The accuracy of a frequency distribution powerlaw fit is sensitive to the choice of dependent (measured) or independent (theoretical) parameter describing the phenomenon, as well as the statistical uncertainty of the number of events in the phenomenon database.

Identifying the underlying physics determining the exact value of the spectral index of the powerlaw, which varies for different size parameters and phenomena, is not yet well understood, but may suggest some form of universality. Since powerlaws are the only statistical distributions that are completely scale-invariant, they offer a unique way to explore the possibility of an underlying universality in nature.

1.3 Self-organized criticality and powerlaw behavior

SOC is also known as the "avalanche concept" and characterizes the behavior of dissipative systems that contain a large number of elements interacting over a short range. The systems evolve to a critical state in which a minor event starts a chain reaction that can affect any number of elements in the system. Such systems are constantly driven by some random energy input evolving into a critical state that is maintained as a powerlaw distribution. SOC theory has the following characteristics:

1. Individual events are statistically independent, spatially and temporally (leading to random waiting time distributions).
2. The size or occurrence frequency distribution is scale-free and can be characterized by a powerlaw function over some size range.
3. The detailed spatial and temporal evolution is complex and involves a fractal geometry and stochastically fluctuating time characteristics (sometimes modelled with 1/f noise, white, pink, red, or black noise).

The concept of SOC evolved from numerical simulations that utilized several relatively simple cellular-automata models. In this context the term cellular refers to the fact that the model is discrete concerning space and the term automaton means that the evolution of the system is self-operating. The traditional SOC cellular automaton model is a regular lattice grid, where redistributions occur with nearest neighbor cells.

In 1987, Per Bak and co-workers presented a model that evolves towards a critical state without any external tuning. This model is often called Per

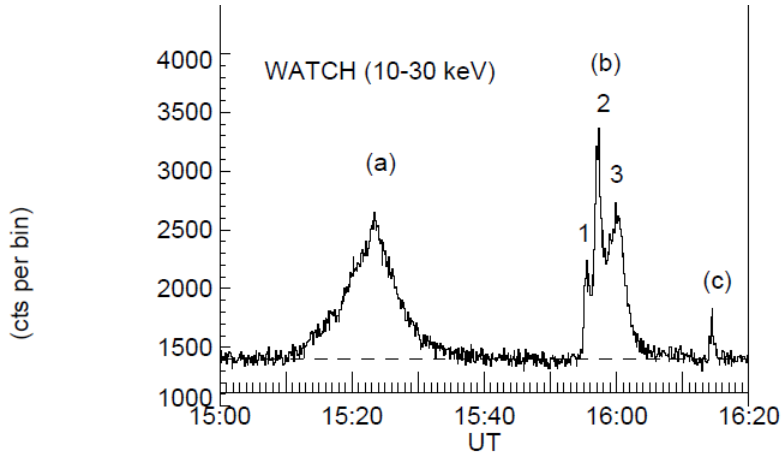


Fig. 1.3: Time profiles of the three WATCH solar flares observed between 15:00 and 16:20 UT on 1990 June 19. The peak count rates are identified as "a", "b" and "c" (Crosby et al. 1998).

Bak's sandpile model or the Bak-Tang-Wiesenfeld (BTW) model. In their first paper (Bak et al. 1987), the BTW model was derived from a model for the dynamics of an array of coupled pendulums. Thereafter the same model was interpreted in terms of sandpile dynamics (Bak et al. 1988). Sandpile avalanches are a paradigm of the SOC theory. The simplicity and beauty of the BTW sandpile model is illustrated in Fig. 1.4.

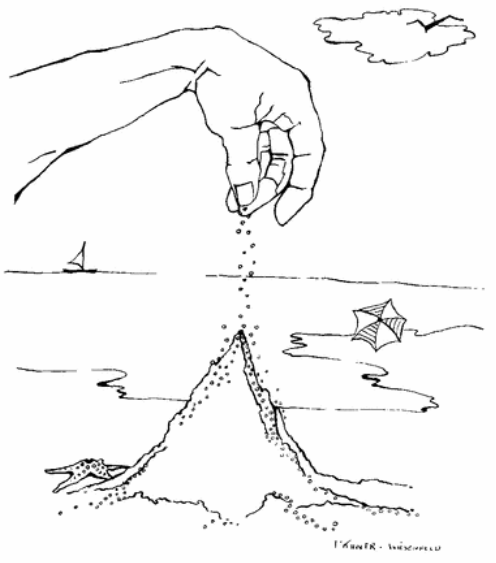


Fig. 1.4: The simplicity and beauty of the Bak-Tang-Wiesenfeld (BTW) sandpile model as drawn by Elaine K. D'Attner Wiesenfeld. Reprinted from Wiesenfeld et al. (1989) with permission.

Another model that exhibits SOC is the forest-fire model (Drossel and Schwabl 1992a; 1992b). In the simplest version of this model, a square grid of sites is considered. At each time step either a tree is planted on a randomly chosen unoccupied site or a spark is dropped on the site. If the spark is dropped on a tree, that tree and all adjacent trees are burned in a model forest-fire. The slider-block model also exhibits SOC and in this model an array of slider blocks are connected to a constant velocity driver plate by driver springs and to each other by connector springs (e.g. Carlson and Langer 1989; Carlson et al. 1994).

1.3.1 Does powerlaw behavior automatically imply SOC?

SOC always involves dynamic "avalanches" with SOC processes occurring spontaneously with an explosive evolution and multiplicative growth via next-neighbor interactions. The restriction to next-neighbor interactions in SOC processes essentially guarantees the statistical independency of individual events. However, there exist related physical processes (turbulence, Brownian motion, percolation, or chaotic systems) that share some of the characteristics of the SOC theory and thus are difficult to discriminate from a SOC process.

For example, self-organization (SO) patterns are quasi-stationary (e.g. geometric patterns in galaxy formation, granulation pattern on the solar surface, ripples on sand dunes) and SO patterns exhibit a close coupling over a large range. If dynamic processes are involved in the formation of SO patterns, they usually involve system-wide processes, such as diffusion, turbulence, convection, magneto-convection, which essentially operate with long-range interactions (via pressure, streams, flows). SO patterns can exhibit scale-free powerlaw distributions of spatial scales (e.g. the Kolmogorov spectrum in turbulent MHD cascades). Powerlaw behavior can therefore not be used as a concise distinction criterion between SO and SOC processes.

There exist also other theoretical models than SOC models that produce powerlaw behavior and many different explanations for the observed powerlaws exist. Examples include Turcotte and Malamud (2004) who related the inverse-cascade model to the results of several cellular automata models and also to real data observed for different natural hazards. Small clusters of (e.g. trees) on a grid coalesce to form larger clusters, and clusters are lost in fires that occur randomly. The result is a self-similar inverse cascade that satisfies an inverse powerlaw distribution of cluster sizes.

Rosner and Vaiana (1978) developed the stochastic relaxation model to describe solar flares. In their model flaring is a stochastic process, energy build-up is exponential between flares, and all the energy built up between flares is released by the following flare whereafter the system returns to its unperturbed "ground state" via the flare. However, their prediction that the duration of energy storage is correlated with flare size was not confirmed by observations (e.g. Lu 1995; Crosby et al. 1998; Wheatland 2000; Georgoulis et al. 2001).

Forced Self-Organized Criticality (FSOC) is an alternative concept that shares all the "avalanche" phenomenology of powerlaw distributions, but is not necessarily self-organized. The key aspect of the FSOC model is that some external dynamics exerts forces on a system to produce powerlaw like distributions of avalanches without internal self-organization (Chang 1999 and references therein).

Turbulence displays many of the common SOC observational signatures, such as the (scale-free) powerlaw distributions of spatial and temporal scales,

the power spectra of time profiles, random waiting time distributions, spatial fractality, and temporal intermittency (e.g. Boffetta et al. 1999). The transition from laminar flow to turbulent flow at a critical Reynolds number represents a similar threshold instability criterion and is the analogy of the critical threshold value used in SOC systems.

Is it possible to distinguish between SOC and turbulence or other processes that show the same scaling? As seen in the above examples powerlaw behavior is not a sufficient argument for SOC. The following three SOC "physics-free" criteria (statistical independence, nonlinear coherent growth, random duration of rise times) have been proposed in Aschwanden (2011):

1. Statistical Independence: events that occur in a SOC system are statistically independent and not causally connected in space or time. Waiting time distributions should be consistent with a stationary or non-stationary Poisson process, in order to guarantee statistical independency by means of probabilities. Time scale separation (the time intervals of the driver are much longer than the time scale of the avalanches, at least for slowly-driven systems).
2. Nonlinear Coherent Growth: time evolution of a SOC event has an initial nonlinear growth phase after exceeding a critical threshold. The nonlinear growth of dissipated energy, or an observed signal that is approximately proportional to the energy dissipation rate, exhibits an exponential-like or multiplicative time profile for coherent processes.
3. Random Duration of Rise Times: if a system is in a state of SOC, the rise time or duration of the coherent growth phase of an avalanche is unpredictable and thus exhibits a random duration. The randomness of rise times can be verified from their statistical distributions being consistent with binomial, Poissonian, or exponential functions.

1.3.2 SOC and SOC-like models

During the last decades new SOC models, as well as non-SOC models have been proposed to explain the powerlaw behavior that is observed in large statistical datasets. These models have gone a step further so as to be able to describe additional characteristics of the observed data (e.g. slope value). For example, Georgoulis and Vlahos (1996) developed a cellular automaton SOC model that simulates flaring activity extending over an active sub-flaring background building on the work by Vlahos et al. (1995). Including both isotropic and anisotropic distributions, as well as a variable magnetic field driving mechanism, they were able to obtain two distinct powerlaw regimes representing possibly two different populations, one associated with standard flares from active regions, and the other with nanoflares from quiet-Sun regions (see Fig. 1.5).

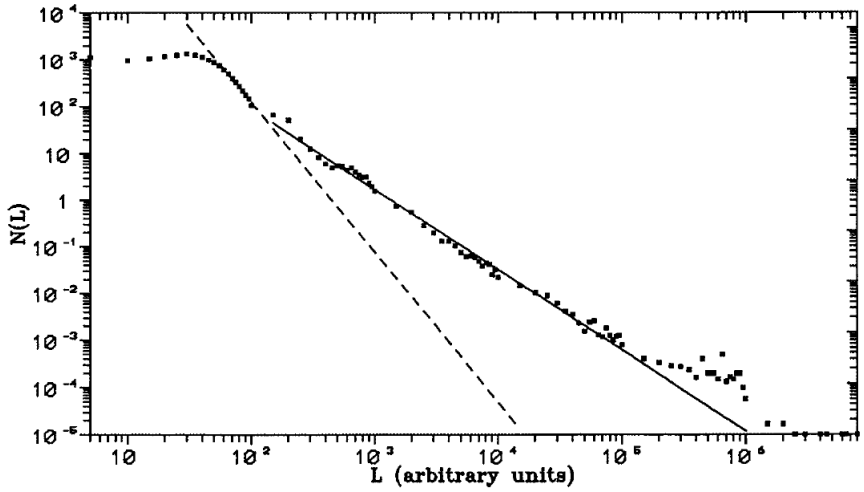


Fig. 1.5: Typical peak-luminosity frequency distribution for a 150 × 150 × 150 grid. The distribution is the average of 10 sample runs. Reprinted from Georgoulis and Vlahos (1996) with permission.

The theoretical side of SOC is covered in the first part of this book. SOC models and SOC related processes from a theoretical point of view are presented by Aschwanden in Chapter 2. He presents how SOC models have built on and been inspired by the original cellular automaton models. Furthermore, a number of alternative dynamical models that are related to SOC models or have similar scaling laws are also presented.

In some respects, chaotic systems exhibit similar complexity as SOC systems, regarding fractality and intermittency; even powerlaw distributions may result in the statistics of chaotic fluctuations. However, the difference to SOC systems is that chaotic systems have these signatures without having an intrinsic mechanism that keeps them near this critical point in a self-organizing way. The strong connections between fractal geometry and SOC from both a mathematical and conceptual understanding are described in Chapter 3 by McAteer.

Percolation controls a transport process that depends on the connectedness and propagation probability of nearest-neighbor elements. It has a lot in common with diffusion, fractal structures, as well as SOC "avalanches". Specifically, it is the fractality and intermittency of the propagating features of the percolation process that are what is in common with a SOC system. In Chapter 4 an introduction to percolation models applied to SOC phenomena is given by Milovanov. SOC-associated phenomena (self-organized turbulence in the Earth's magnetotail, phase transitions in SOC systems, mixed SOC-coherent behavior, periodic and auto-oscillatory patterns of behavior), are also discussed.

In probability theory, a branching process expresses next-neighbor interactions in terms of probabilities, a concept similar to how redistribution operates in cellular automaton models. Litvinenko (1998) applies methods of the branching theory to the Macpherson and MacKinnon (1997) cellular automation model for the occurrence of solar flares. In Chapter 5 Corral and Font-Clos present how models based on the branching process are applied to "natural hazards". Applied to earthquakes a branching process implies the activation or slip of a fault segment that can trigger other segments to slip, with a certain probability, and so on.

In contrast to the SOC cellular automaton models that were described in the beginning of Section 3, networks are irregular nets of nodes that are interconnected in manifold patterns, containing nearest-neighbor connections and in some cases also arbitrary non-local, long-range connections. Zou, Heitzig, Small, and Kurths present in Chapter 6 recurrence networks as a novel tool of nonlinear time series analysis allowing the characterization of higher-order geometric properties of complex dynamical systems based on recurrences in phase space, which are a fundamental concept in classical mechanics. The main part of the Chapter is based on the Zou et al. (2012) paper. They demonstrate that recurrence networks obtained from various deterministic model systems as well as experimental data naturally display powerlaw degree distributions with scaling exponents that can be derived exclusively from the systems' invariant densities.

SOC models are said to be slowly driven interaction dominated threshold systems (Jensen 1998). The scaling behavior of a SOC model can be related to some underlying continuous phase transition, which is triggered by the system self-organizing to the critical point. Pruessner presents in Chapter 7 an overview of how computer simulations are used to reproduce SOC behavior. The Chapter presents in more detail the original SOC models and shows how the observed scaling behavior can be better quantified. Furthermore, detailed functions that are used in SOC algorithms are described.

1.4 Where is SOC observed ?

Following Bak's above mentioned pioneering work in 1987, not only was there an "avalanche" in SOC studies from the modelling side, but people started to search for signatures of SOC everywhere in natural occurring phenomena from the Sun to the Earth. In the second part of this book the Chapters concern how SOC is observed in all types of natural phenomena.

1.4.1 Phenomena on Earth showing SOC behavior

Back in the early 1990s simple sandpile experiments were performed to reproduce SOC (Held et al. 1990; Bak and Chen 1991). In simple terms sand grains are added to a sandpile until the slope of the sandpile reaches a critical value and an avalanche occurs. Later ricepile experiments were also performed (e.g. Frette et al. 1996). In Chapter 8 Pruessner describes many of the SOC laboratory experiments that have been performed during these last decades.

In parallel to experiments in the laboratory numerous studies have been performed on naturally occurring phenomena on Earth. The most well-known Earth-based SOC phenomenon that has been and is being studied is the earthquake. They are triggered when a mechanical instability occurs and a fracture (the sudden slip of a fault) appears in a part of the Earth's crust. Earthquakes are associated with the slider-block model (e.g. Carlson and Langer 1989; Carlson et al. 1994).

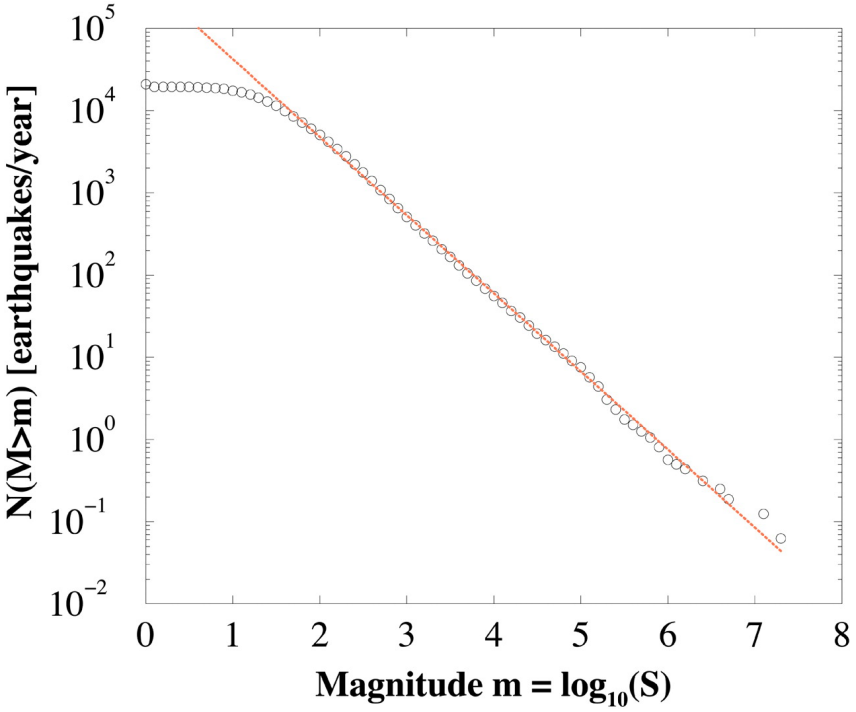


Fig. 1.6: Earthquake magnitude distribution showing a powerlaw behavior over six decades. The graph follows $log10N(M > m)\alpha - bm$, where b is the Gutenberg-Richter exponent $b = 1$ (dashed red line has a slope value of -0.95). Reprinted from Christensen et al. (2002) with permission.

Fig. 1.6 shows the earthquake magnitude distribution for the Southern California region, which is the number of earthquakes per year with magnitude (M > m). The dashed red line has a slope value of -0.95 and shows the Gutenberg-Richter Law with a gradient b $\approx$ 1 (Christensen et al. 2002; Christensen and Moloney 2005). As was seen for the solar flare distribution (Fig. 3.1) there is a turn-over in the lower end of the distribution due to problems associated with detecting small earthquakes. In Chapter 9 Sachs et al. present SOC applied to complex earthquakes and how this concept is observed in data and models, as well as how it is applied to forecasting. From a mitigation perspective the main question to answer is when will an earthquake occur? A new type of forecast based on the Natural Time Weibull (NTW) model shows promising results (Rundle et al. 2012) and is based on the idea of "filling in" a fat-tailed (scaling) distribution.

Besides seismology and earthquakes SOC behavior has been found in a wide range of geophysical systems. Forest-fires occurring on Earth are associated with the forest-fire model that also exhibits SOC behavior (Drossel and Schwabl 1992a; 1992b). In Chapter 10 Hergarten presents the forest-fire model and actual wildfires. Specifically he presents studies performed on observed wildfires and how the forest-fire model is able to reproduce the difference between natural and man-made forest-fires.

Natural landslide events are commonly associated with a trigger (e.g. earthquake (minutes after), a rapid snowmelt (hours to days), or an intense rainfall (days to weeks)) and range in size from a single landslide to many thousands. A landslide is a natural phenomenon associated with the Per Bak sandpile model. The impact of landslides is often limited to smaller areas than for example the damage caused by earthquakes. However, there is growing evidence for powerlaw size distributions in different types of landslides. In Chapter 11 Hergarten presents how SOC behavior is observed in landslides. One of the main conclusions is that scaling exponents found for regolith landslides strongly differ from those found for rockfalls and rockslides, but each of the classes may be characterized by a universal scaling exponent.

1.4.2 Phenomena in space showing SOC behavior

Besides the powerlaw behavior that is observed in natural occurring phenomena on Earth, it is also observed in a wide range of phenomena occurring in space. Like their counterparts on Earth they too come in all sizes and durations. Our closest star, the Sun, drives the continuous changing conditions in the space environment - the local space weather. It is well known that the solar corona is a very dynamic region which is the source of many phenomena (e.g. solar flares, coronal mass ejections, solar energetic particle events). It was therefore not surprising that it was on the Sun where SOC signatures were first discovered in space.

Like earthquakes are to SOC on Earth, solar flares are to SOC in space. Solar flares (including micro- and nano-flares) are one of the most compelling examples of SOC type behavior; their powerlaw distribution covers over eight orders of magnitude (Aschwanden 2011). Crosby et al. (1993) presents a summary of 13 powerlaw slopes (12 originating from observational results and 2 from modelling) found for different solar related parameters that were published over two decades ago. Since then this list has grown exponentially; see Aschwanden (2011) for an extensive overview. In Chapter 12 Charbonneau presents SOC behavior observed in solar phenomena as well as SOC models developed to describe the distribution of solar flares.

From the dawn of the space era observations in space have provided convincing evidence indicating that certain space plasma processes are in states of complexity and SOC, especially with the above discovery of the apparent powerlaw probability distribution of solar flare intensities. The most studied extra-solar SOC phenomena are stellar flares (Audard et al. 2000) and accretion or black-hole objects (Negoro et al. 1995; Mineshige and Negoro 1999). In Chapter 13 Aschwanden gives an overview of how SOC behavior has been observed in astrophysics. He compares theoretical predictions based on the fractal-diffusive self-organized criticality (FD-SOC) model (Aschwanden 2012) with observed powerlaws of size distributions observed in astrophysical systems. It is found that the generalized FD-SOC model can explain a large number of astrophysical observations (e.g. lunar craters, asteroid belts, Saturn rings, outer electron radiation belt enhancements, solar flares, soft gamma-ray repeaters, and blazars) and discriminate between different scaling laws of astrophysical observables.

1.5 Searching for a common signature: What does it all mean?

Does there exist a common "avalanche" signature? If yes, what does this all mean in a global way? How does the world and the Universe connect? Is there a universal signature that places constraints on the energy distribution of everything in the Universe? SOC behavior provides us a unique way to interpret the behavior of "avalanches" in a global way - is there a common thread in nature?

Producing frequency distributions on observational "avalanche" data displays powerlaw behavior in most instances. Each type of phenomenon is observed to have a range of powerlaw slope values. For example, considering solar flare hard X-ray parameters, in general the powerlaw slope ranges between -1.4 through -2.4 (Crosby 2011). The difference in value is also observed on measured parameters of the same type of "avalanche" suggesting that the slope may be detector dependent. An independent parameter based on theory is therefore better for comparison purposes. The energy released in an

"avalanche" is such an independent parameter and can be compared directly without considering the measured data. In this way it can be investigated if energy is released in some type of universal way. For example, powerlaw frequency distributions of the energy released in some natural phenomena (e.g. solar flares, transient brightenings, nanoflares, ionospheric emissions "auroral blobs", earthquakes) are found to be similar (slope value of the powerlaw is approximately -1.5), see Crosby (2011).

In summary, modelling a given type of "avalanche" as a complex system in a self-organized critical state provides a good context to understand the frequency distributions of the parameters describing the phenomenon. It may be SOC or another theoretical model that can account for the observed powerlaw behavior, or it may end up being a mixture of many physical processes occurring simultaneously. However, as pointed out by Crosby (2011) in a SOC review paper, whatever the theory, the observations themselves provides one with useful information when performing frequency distributions on the data. Table 4.2 summarizes the information one obtains by performing frequency distributions on "avalanches" both on data obtained by observations and on SOC model outputs. It should also be emphasized that any model, be it SOC or non-SOC, must be able to reproduce what the observations are showing (model validation).

Performing frequency distributions on "avalanches" is also a tool for practical applications in regard to mitigation and risk analysis. Indeed, the applicability of powerlaw statistics to natural hazards has important practical implications. In the paper by Sachs et al. (2012) the question "Are dragon-kings relevant to probabilistic hazard assessment?" is addressed by studying earthquakes, volcanos, etc. Sornette (2009) developed the concept of the unexpected "dragon-kings" to describe this class of extreme events that are significantly larger than the extrapolation of the powerlaw scaling of their smaller counterparts. Probabilistic seismic hazard studies often extrapolate the rate of occurrence of small earthquakes to quantify the probability of occurrence of large earthquakes. Sachs et al. (2012) argue that this extrapolation would therefore not be valid if "dragon-king" earthquakes occur and one of their examples (cumulative frequency magnitude distribution of earthquakes in the Parkfield aftershock region covering 1972 to 2009) is displayed in Fig. 1.7. They find that the Parkfield main shock with m = 5.95 (identified as the red star on Fig. 1.7) is found to lie above the extrapolation of the powerlaw correlation of the smaller earthquakes. The authors question whether the difference between m = 5.65 based on the extrapolation and m = 5.95 can be attributed to the statistical variability of the characteristic earthquakes or if this is indeed a "dragon-king" type earthquake.

Fig. 1.8 shows the cumulative number of volcanic eruptions (Nc) during the period 1800-2002 with dense rock equivalent volume (V_{DRE}) greater than V_{DRE} as a function of V_{DRE}. Sachs et al. (2012) find that global frequency size distributions of earthquakes and volcanic eruptions exhibit powerlaw behavior for small sizes but a roll over for large events similar to the behavior

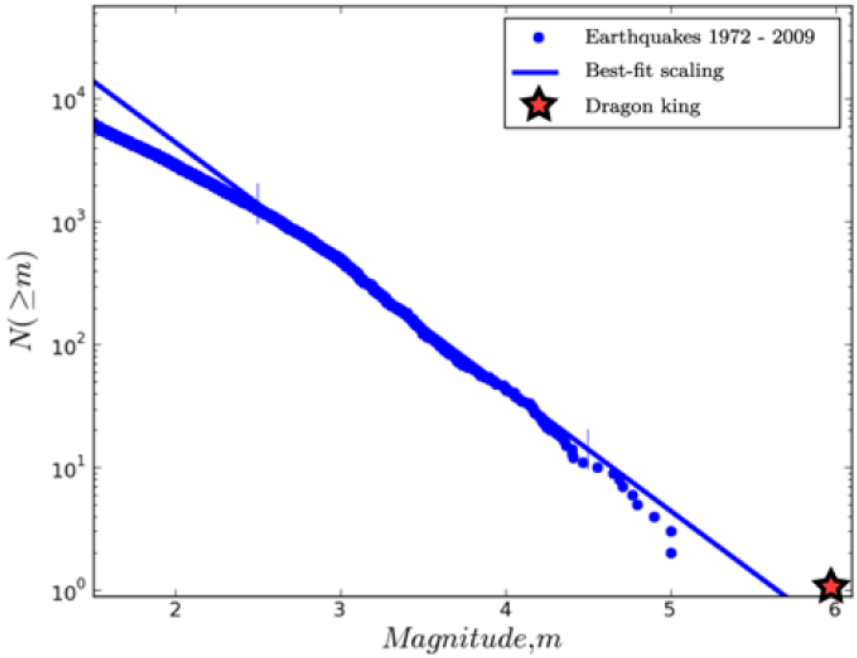


Fig. 1.7: Cumulative number of earthquakes with magnitude greater than m as a function of m for the Parkfield earthquake cycle 1972 to 2009. The best-fit scaling is shown as the blue line. The m = 5.95 Parkfield earthquake is shown as a "dragon-king" (identified as the red star). Reprinted from Sachs et al. (2012) with permission.

of the forestfire model for small firing frequencies. The Lake Toba Sumatra volcano is identified as the red star on Fig. 1.8 and is estimated to have erupted $V_{DRE} = 2{,}750 \pm 250 km^3$ of dense rock equivalent 73,500 ± 500 years ago. However, the largest eruption expected in 73,500 years, extrapolating the powerlaw on Fig. 1.8, would have yielded a volume $V_{DRE} = 7.9 \times 10^6$ km^3. This is statistically not consistent and Sachs et al. (2012) therefore suggest a powerlaw with an exponential roll-over for volumes greater then $V_{DRE} \approx 10^2$ km^3 as the most realistic fit for the distribution displayed in Fig. 1.8. Such results show the important practical implications that powerlaw statistics has in regard to natural hazards and question whether a powerlaw is always the best representation for the very large "black swans" events. However, this exponential roll-over effect may also be due to observations having not been performed over a long enough period of time to cover all the statistics of the "avalanches" as mentioned in Section 2.

During the last years there has been renewed interest in SOC and its potential applications both from theoretical and practical point-of-views. Several

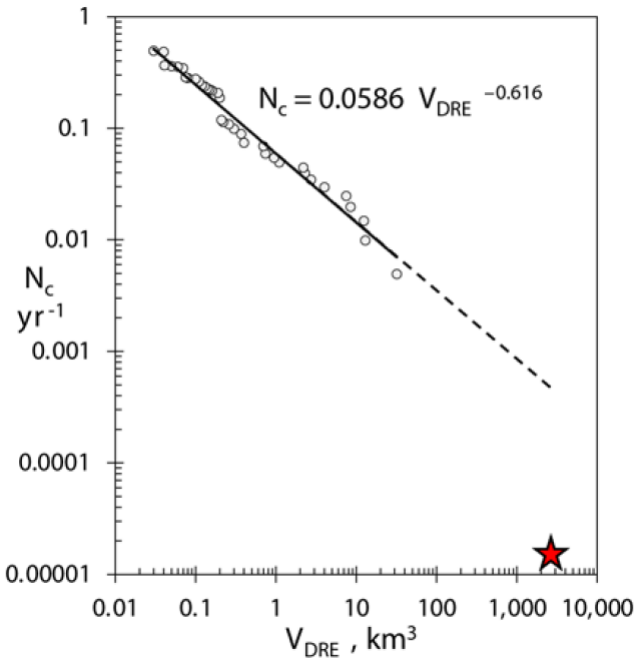


Fig. 1.8: Cumulative number of volcanic eruptions (Nc) during the period 1800-2002 with dense rock equivalent volume (V_{DRE}) greater than V_{DRE} as a function of V_{DRE}. The best-fit powerlaw scaling is also shown along with the Toba eruption in Sumatra (identified as the red star) occurring 73,500 $\pm$ 500 years ago. Reprinted from Sachs et al. (2012) with permission.

cross-disciplinary initiatives have currently been taken such as the International Space Science Institute International Team entitled "Self-Organized Criticality and Turbulence" that is made up of an interdisciplinary group of team members covering both the space- and Earth-sciences. Its aim is to cross-compare observations, to discuss SOC, SOC-related (such as turbulence), and non-SOC theoretical models, and to establish a diagnostic metrics between observations and theoretical models that yield new physical insights into SOC phenomena and complexity in nature.

As was shown in this Chapter SOC behavior has become one way of interpreting the powerlaw behavior observed in natural occurring phenomena on the Sun down to the Earth. This book, based on an inter-disciplinary approach, presents the many sides of SOC both from the theoretical side of the story as well as the observational.

Acknowledgements The author would like to thank the editor of this SOC book for asking her to write this Introduction Chapter. The author acknowledges the Inter-

Table 1.1: What do frequency distributions teach us? Based on Crosby (2011).

Observational statistics	- Frequency distributions performed on datasets describing natural dynamical phenomena (e.g. solar flares, earthquakes) exhibit powerlaw behavior.
Measurement problems and biases	- Turn-overs in the lower end of the frequency distribution for various phenomena may be attributed to detector sensitivity (missing the small events in the background noise). - Exponential turn-overs in the upper end of the frequency distribution for various phenomena may be due to either the length of the dataset (missing of long-term statistics) or finite-size effects of the system. - Measured parameters are detector dependent and may give bias in the slopes of the frequency distributions for comparison purposes.
Numerical SOC observations	-Various concepts/models exist that produce powerlaw behavior such as SOC. -SOC models are able to reproduce the results found when performing frequency distributions on measured data.
Statistical commonalities in SOC statistics	-Powerlaw behavior is found to be a universal characteristic defining natural dynamic phenomena (e.g. solar flares, earthquakes). -For each type of phenomenon most distributions performed on observational data can be represented by powerlaws having a range of powerlaw slope values. -Frequency distributions of the energy released in solar flares, transient brightenings, nanoflares, ionospheric emissions and earthquakes are found to be similar (slope value of the powerlaw is approximate -1.5).
Interpretations of physical processes	-Powerlaw frequency distributions result from nonlinear or coherent processes, have no characteristic spatial scale and are the hallmark of nonlinear dissipative systems. -Powerlaw frequency distributions of the energy released in some natural phenomena are found to be similar (may suggest that energy is released in some type of universal way).
Mitigation and risk analysis	-Results from frequency distributions provide limits to the maximum strength of a phenomenon, vital for mitigation studies - probability of extreme events occurring (limit to the size of an event over a given time period). -Implementing frequency distributions into the engineering approach "empirical models" is useful for design studies as well as probabilistic hazard assessment.

national Space Science Institute "Self-Organized Criticality and Turbulence" International Team of which she is a team member and which has been an inspiration for the

writing of this Introduction Chapter. She also would like to acknowledge the reviewer who took the time to provide useful feed-back thereby improving the objective of this Introduction Chapter.

References

Aschwanden, M.J.: Self-Organized Criticality in Astrophysics. The Statistics of Nonlinear Processes in the Universe, ISBN 978-3-642-15000-5, 416p. Springer-Praxis: New York (2011)

Aschwanden, M.J.: A statistical fractal-diffusive avalanche model of a slowly-driven self-organized criticality system. Astron. Astrophys. **539**, A2 (2012)

Audard, M., GĂĽdel, M., Drake, J.J., and Kashyap, V.L.: Extreme-ultraviolet flare activity in late-type stars. Astrophys. J. **541**, 396–409 (2000)

Bak, P.: How Nature works. Springer-Verlag: New York (1996)

Bak, P., Tang, C. and Wiesenfeld, K.: Self-organized criticality - An explanation of 1/f noise. Physical Review Lett. **59/27**, 381–384 (1987)

Bak, P., Tang, C. and Wiesenfeld, K.: Self-organized criticality. Physical Review A. **38/1**, 364–374 (1988)

Bak, P. and Chen, K.: Self-organized criticality. Scientific American **264**, 46–53 (1991)

Boffetta, G., Carbone, V., Giuliani, P., Veltri, P. and Vulpiani, A.: Power laws in solar flares: self-organized criticality or turbulence. Phys. Rev. Lett. **83(2)**, 4662–4665 (1999)

Carlson, J.M. and Langer, J.S.: Properties of Earthquakes Generated by Fault Dynamics. Phys. Rev. Lett. **62**, 2632–2635 (1989)

Carlson, J.M., Langer, J.S. and Shaw, B.E.: Dynamics of earthquake faults. Rev. Modern Phys. **66**, 657–670 (1994)

Chapman, S.C., Watkins, N.W., Dendy, R.O., Helander, P. and Rowlands., G.: A simple avalanche model as an analogue for magnetospheric activity. Geophys. Res. Lett. **25/13**, 2397–2400 (1998)

Chang, T.S.: Self-organized criticality, multi-fractal spectra, sporadic localized reconnections and intermittent turbulence in the magnetotail. Phys. Plasmas. **6/11**, 4137–4145 (1999)

Charbonneau, P., McIntosh, S.W., Liu, H.L. and Bogdan, T.J.: Avalanche models for solar flares. Solar Phys. **203**, 321–353 (2001)

Christensen, K., Danon, L., Scanlon, T. and Bak, P.: Unified scaling law for earthquakes. PNAS. **99**, suppl. 1, 2509–2513 (2002)

Christensen, K. and Moloney, N.R.: Complexity And Criticality, 392p. Imperial College Press, London, U.K. (2005)

Crosby, N.B., Aschwanden, M.J. and Dennis, B.R.: Frequency Distributions and Correlations of Solar X-Ray Flare Parameters. Solar Phys. **143**, 275–299 (1993)

Crosby, N., Vilmer, N., Lund, N. and Sunyaev, R.: Deka-keV X-Ray Observations of Solar Bursts with WATCH/GRANAT: frequency distributions of burst parameters. Astronom. Astrophys. **334**, 299–313 (1998)

Crosby, N.B.: Frequency distributions: From the Sun to the Earth. Nonlinear Processes in Geophysics. **18/6**, 791–805 (2011)

Drossel, B. and Schwabl, F.: Self-organized critical forest-fire model. Phys. Rev. Lett. **69**, 11, 1629–1632 (1992a)

Drossel, B. and Schwabl, F.: Self-organized criticality in a forest-fire model. Physica A. **191**, 1-4 47–50 (1992b)

Frette, V., Christensen, K., Malthe-Sorenssen, A., Feder, J., Jessang, T. and Meakin, P.: Avalanche dynamics in a pile of rice. Nature. **379**, 49–52 (1996)

Georgoulis, M. and Vlahos L.: Coronal Heating by Nanoflares and the Variability of the Occurence Frequency in Solar Flares. Astrophys. J. Letters. **469**, L135–L138 (1996)

Georgoulis, M., Vilmer, N. and Crosby, N.: A Comparison between statistical properties of solar X-ray flares and avalanche predictions in cellular automata statistical flare models, A&A. **367**, 326–338 (2001)

Held, G.A., Solina, D.H., Solina, H., Keane, D.T., Haag, W.J., Horn, P.M. and Grinstein, G: Experimental Study of Critical mass Fluctuations in an Evolving Sandpile. Phys. Rev. Lett. **65**, 9, 1120–1123 (1990)

Hergarten, S.: Self-Organized Criticality in Earth Systems, 282p. Springer-Verlag: Germany (2002)

Jensen, H.J.:Emergent complex behavior in physical and biological systems, Cambridge University Press (1998)

Litvinenko, Y.E.: Analytical results in a cellular automaton model of solar flare occurrence. Astronomy and Astrophysics. **339**, L57–L60 (1998)

Lu, E.T.: Constraints on Energy Storage and Release Models for Astrophysical Transients and Solar Flares. Astrophys. J. **447**, 416–418 (1995)

Macpherson, K.P., and MacKinnon, A.L.: One-dimensional percolation models of transient phenomena. Physica A. **243**, 1–13 (1997)

Mineshige, S. and Negoro, H.: Accretion disks in the context of self-organized criticality: How to produce 1/f fluctuations? ASP Conf. Ser. **161**, 113–128 (1999)

Negoro H., Kitamoto S., Takeuchi M. and Mineshige S.: Statistics of X-ray fluctuations from Cygnus X-1: Reservoirs in the disk? Astrophys. J. **452**, L49–L52 (1995)

Pruessner, G.: Self-organised criticality, Cambridge University Press: Cambridge, UK (2012)

Rosner, R. and Vaiana, G.S.: Cosmic flare transients â“ Constraints upon models for energy storage and release derived from the event frequency distribution. The Astrophysical Journal. **222**, 1104–1108 (1978)

Rundle, J.B., Holliday, J.R., Graves, W., Turcotte, D.L., Tiampo, K.F. and Klein W.: Probabilities for large events in driven threshold systems. Phys. Rev. E. **86**, 021106 (2012)

Sachs, M.K., Yoder, M.R., Turcotte, D.L., Rundle, J.B. and Malamud, B.D.: Black swans, power laws, and dragon-kings: Earthquakes, volcanic eruptions, landslides, wildfires, floods, and SOC models. Eur. Phys. J. Special Topics **205**, 167–182 (2012)

Sornette, D.: Critical phenomena in natural sciences: chaos, fractals, self-organization and disorder: concepts and tools, 528p. Springer: Heidelberg (2004)

Sornette, D.: Black Swans and the Prediction of Crises. International Journal of Terraspace Science and Engineering 2 **1**, 1–18 (2009)

Taleb, N.N.: The Black Swan: The Impact of the Highly Improbable. 480p. Random House: New York (2007)

Turcotte, D.L.: Self-organized criticality. Rep. Prog. Phys. **62**, 1377–1429 (1999)

Turcotte, D.L. and Malamud, B.D.: Landslides, Forest Fires, and Earthquakes: Examples of Self-Organized Critical Behavior. Physica A. **340**, 580–589 (2004)

Vlahos, L., Georgoulis, M., Kluiving, R. and Paschos, P.: The statistical flare. Astronomy and Astrophysics. **299**, 897–911 (1995)

Wheatland, M.S.: Do solar flares exhibit an interval-size relationship? Solar Phys. **191**, 381–389 (2000)

Wiesenfeld, K., Tang, C., Bak, P.: A Physicist's Sandbox. Journal of Statistical Physics, **54**, 1441–1458 (1989)

Willinger, W., Alderson, D., Doyle, J.C. and Li, L.: More "normal" than normal: scaling distributions and complex systems. in *Proceedings of the 2004 Winter Simulation Conference*, ed. by R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters. (IEEE Press, Piscataway, NJ), 130–141 (2004)

Zou, Y., Heitzig, J., Donner, R.V., Donges, J.F., Farmer, J.D., Meucci, R., Euzzor, S., Marwan, N. and Kurths, J.: Power-laws in recurrence networks from dynamical systems. EPL (Europhysics Letters). **98**, (2012)

Chapter 2
Theoretical Models of SOC Systems

Markus J. Aschwanden

How can the universe start with a few types of elementary particles at the big bang, and end up with life, history, economics, and literature? The question is screaming out to be answered but it is seldom even asked. Why did the big bang not form a simple gas of particles, or condense into one big crystal? (Bak 1996). The answer to this fundamental question lies in the tendency of the universal evolution towards complexity, which is a property of many nonlinear energy dissipation processes. Dissipative nonlinear systems generally have a source of free energy, which can be partially dissipated whenever an instability occurs. This triggers an avalanche-like energy dissipation event above some threshold level. Such nonlinear processes are observed in astrophysics, magnetospheric physics, geophysics, material sciences, physical laboratories, human activities (stock market, city sizes, internet, brain activity), and in natural hazards and catastrophes (earthquakes, snow avalanches, forest fires). A tentative list of *self-organizing criticality (SOC)* phenomena with the relevant sources of free energy, the physical driver mechanisms, and instabilities that trigger a SOC event are listed in Table 2.1.

A prominent theory that explains such nonlinear energy dissipation events is the so-called *Self-organized criticality (SOC)* concept, first pioneered by Bak et al. (1987, 1988) and Katz (1986), and simulated with cellular automaton or other lattice models, which mimic nearest-neighbour interactions leading to complex patterns. The topic of SOC is reviewed in recent reviews, textbooks, and monographs (e.g., Bak 1996; Jensen 1998; Turcotte 1999; Charbonneau et al. 2001; Hergarten 2002; Sornette 2004; Aschwanden 2011a; Crosby 2011; Pruessner 2012). A disclaimer has to be made, that we discuss in this chapter the basic ideas of SOC only, mostly with the special

Markus J. Aschwanden, Solar and Astrophysics Laboratory (LMSAL), Lockheed Martin, Advanced Technology Center, 3251 Hanover St., Palo Alto, CA 94304, USA; e-mail: aschwandenlmsal.com.

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

application to astrophysics in mind, while an extensive set of other analytical and numerical SOC models can be found elsewhere (e.g., Pruessner 2012).

SOC can be considered as a basic physics phenomenon - universally occurring in systems with many coupled degrees of freedom in the limit of infinitesimal external forcing. This theory assumes a critical state that is robust in the sense that it is self-organizing, like a critical slope of a sandpile is maintained under the steady (but random) dropping of new sand grains on top of the pile. Individual avalanches occur with sizes that are broadly distributed and many orders of magnitude larger than the initial perturbation. Sandpile avalanches are a paradigm of the SOC theory, which has the following characteristics in the slowly-driven limit: (1) Individual events are triggered independently of each other in space and time (leading to random waiting time distributions); (2) The size or occurrence frequency distribution is scale-free and can be characterized by a powerlaw function over some size range; (3) The detailed spatial and temporal evolution is complex and involves a fractal geometry and stochastically fluctuating (intermittent) time characteristics (sometimes modeled with 1/f-noise, white, pink, red, or black noise).

There are some related physical processes that share some of these characteristics, and thus are difficult to discriminate from a SOC process, such as turbulence, Brownian motion, percolation, or chaotic systems (Fig. 2.1, right column). A universal SOC theory that makes quantitative predictions of the powerlaw-like occurrence frequency and waiting time distributions is still lacking. We expect that the analysis of large new databases of space observations and geophysics records (available over at least a half century now) will provide unprecedented statistics of SOC observables (Fig. 2.1, left column), which will constrain the observed and theoretically predicted statistical distribution functions (Fig. 2.1, middle column), and this way confirm or disprove the theoretical predictions of existing SOC theories and models (Fig. 2.1, right column). Ultimately we expect to find a set of observables that allows us to discriminate among different SOC models as well as against other nonlinear dissipative processes (e.g., MHD turbulence). In the following we will describe existing SOC models and SOC-related processes from a theoretical point of view.

2.1 Cellular Automaton Models (CA-SOC)

The theoretical models can be subdivided into: (i) a mathematical part that deals with the statistical aspects of complexity, which is universal to all SOC phenomena and essentially "physics-free" (such as the powerlaw-like distribution functions and fractal dimensions), and (ii) a physical part that links the SOC avalanche volume to a physical observable in terms of a particu-

Table 2.1: Examples of physical processes with SOC behavior (for references see Table 13.2 in chapter 13 on *SOC systems in astrophysics* in this book).

SOC Phenomenon	Source of free energy or physical mechanism	Instability or trigger of SOC event
Galaxy formation	gravity, rotation	density fluctuations
Star formation	gravity, rotation	gravitational collapse
Blazars	gravity, magnetic field	relativistic jets
Soft gamma ray repeaters	magnetic field	star crust fractures
Pulsar glitches	rotation	Magnus force
Blackhole objects	gravity, rotation	accretion disk instability
Cosmic rays	magnetic field, shocks	particle acceleration
Solar/stellar dynamo	magnetofriction in tachocline	magnetic buoyancy
Solar/stellar flares	magnetic stressing	magnetic reconnection
Nuclear burning	atomic energy	chain reaction
Saturn rings	kinetic energy	collisions
Asteroid belt	kinetic energy	collisions
Lunar craters	lunar gravity	meteroid impact
Magnetospheric substorms	electric currents, solar wind	magnetic reconnection
Earthquakes	continental drift	tectonic slipping
Snow avalanches	gravity	temperature increase
Sandpile avalanches	gravity	super-critical slope
Forest fires	heat capacity of wood	lightening, campfire
Lightenings	electrostatic potential	discharge
Traffic collisions	kinetic energy of cars	driver distraction, ice
Stockmarket crashes	economic capital, profit	political event, speculation
Lottery wins	optimistic buyers	random drawing system

lar physical mechanism (for instance in form of a scaling law between the avalanche volume and the observed emission).

2.1.1 Statistical Aspects

The original concept of self-organized criticality was pioneered by Bak et al. (1987, 1988) and Katz (1986), who used the paradigm of a sandpile with a critical slope to qualitatively illustrate the principle of self-organized criticality. The first theoretical description of the SOC behavior of a nonlinear system was then studied with a large number of coupled pendulums and with numerical experiments that have been dubbed *cellular automaton* models (CA-SOC). In essence, a SOC system has a large number of elements with coupled degrees of freedom, where a random disturbance causes a complex

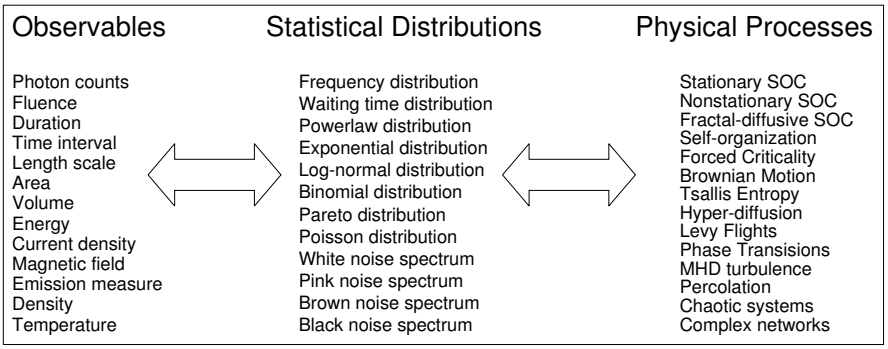


Fig. 2.1: Matrix of observables, statistical distributions, and physical models that need to be defined in order to discriminate SOC from non-SOC processes.

dynamical spatio-temporal pattern. Although such a nonlinear system obeys classical or quantum-mechanical physics, it is impractical to describe it analytically because of the large number of degrees of freedom. The situation is similar to the N-body problem, which cannot be solved analytically for more complex configurations than a two-body system. However, such SOC systems can easily be simulated with a numerical computer program, which starts with an initial condition defined in a lattice grid, and then iteratively updates the dynamical state of each system node, and this way mimics the dynamical evolution and statistical distribution functions of SOC parameters. Hence, such numerical lattice simulations of SOC behavior (i.e., cellular automatons) were the first viable methods to study the SOC phenomenon on a theoretical basis.

What are the ingredients and physical parameters of a cellular automaton model? What is common to all cellular automaton models of the type of Bak, Tang, and Wiesenfeld (1987, 1988), briefly called BTW models, is: (1) A S-dimensional rectangular lattice grid, say a coordinate system $x_{i,j,k}, i = 1, ..., n; j = 1, ..., n, k = 1, ..., n$ for a 3-dimensional case $S = 3$ with grid size n; (2) a place-holder for a physical quantity $z_{i,j,k}$ associated with each cellular node $x_{i,j,k}$, (3) a definition of a critical threshold $z_{\rm crit}$, (4) a random input $\Delta z_{i,j,k}$ in space and time; (5) a mathematical re-distribution rule that is applied when a local physical quantity exceeds the critical threshold value, for instance a critical slope of a sandpile, which adjusts the state of the nearest-neighbor cells, i.e.,

$$\begin{array}{ll} z_{i,j,k} \mapsto z_{i,j,k} + 1 & \text{initial input} \\ z_{i,j,k} \mapsto z_{i,j,k} - 8 & \text{if } z_{i,j,k} \geq 8, \\ z_{i\pm1,j\pm1,k\pm1} \mapsto z_{i\pm1,j\pm1,k\pm1} + 1 & \end{array} \tag{2.1}$$

for the eight nearest neighbors in a 3D-lattice grid, and (6) iterative steps in time to update the system state $z_{i,j,k}(t)$ as a function of time t. Such cellular automaton simulations usually start with a stable initial condition (e.g., an empty system with $z_{i,j,k}(t = 0) = 0$), and have then to run several million time steps before they reach a critical state. Once they reach the critical state, avalanches of arbitrary sizes can be triggered by the random input of an infinitesimal disturbance $\Delta z_{i,j,k}(t)$, separated by quiescent time intervals in between, in the case of slow driving. Statistics of the occurrence frequency distributions of avalanche sizes or durations yields then the ubiquitous powerlaw-like distribution functions. Fig. 2.2 shows an example of a 2D-lattice simulation, displaying a snapshot of a large avalanche (top left panel) as well as the statistical distribution function of avalanche sizes L (left) and durations T (right panels).

Thus, such SOC cellular automaton models primarily generate the spatio-temporal dynamics of complex systems and the related statistics of emerging spatio-temporal patterns. They can simulate the distribution function $N(L)dL$ of the size or length scale L of dynamical events (also called "avalanches"), or the time scale T of an avalanche event. Secondary dynamical parameters can be derived, such as the instantaneous avalanche volume $V(t)$, or the total time-integrated volume of the avalanche $V(t < T)$. However, such SOC cellular automaton models are "physics-free", because the dynamics of an event is described by a mathematical rule that substitutes for the unknown physical process, but is thought to be an universal property of SOC systems in a critical state. Applying cellular automaton simulations to observed phenomena requires than to substitute a physical quantity into the "place-holder" variable $z_{i,j,k}$, which is otherwise without a physical interpretation in cellular automata, except for defining the instability criterion $z_{i,j,k} > z_{\rm crit}$ in terms of a critical threshold $z_{\rm crit}$.

There is a large industry of cellular automaton models, mostly modified versions of the original BTW-model, such as the lattice-gas model (Jensen 1998), Conway's game of life model (Bak, Chen and Creutz 1989), traffic jam simulations (Nagel and Paczuski 1995), multiple-strategy agent-based models applied to the financial market (Feigenbaum 2003), punctuated equilibrium models applied to biophysics (Bak and Sneppen 1993), slider-block spring models applied in geophysics (Burridge and Knopoff 1967), forest-fire models (Malamud et al. 1998; reviewed in Turcotte 1999), applications to magnetospheric substorms (Takalo et al. 1993; 1999a,b), to solar flares (Lu and Hamilton 1991; Charbonneau et al. 2001), and to stellar accretion disks (Mineshige et al. 1994a,b; Pavlidou et al. 2001). More complete reviews of such cellular automaton models are given in Turcotte (1999), Aschwanden (2011a), and Pruessner (2012).

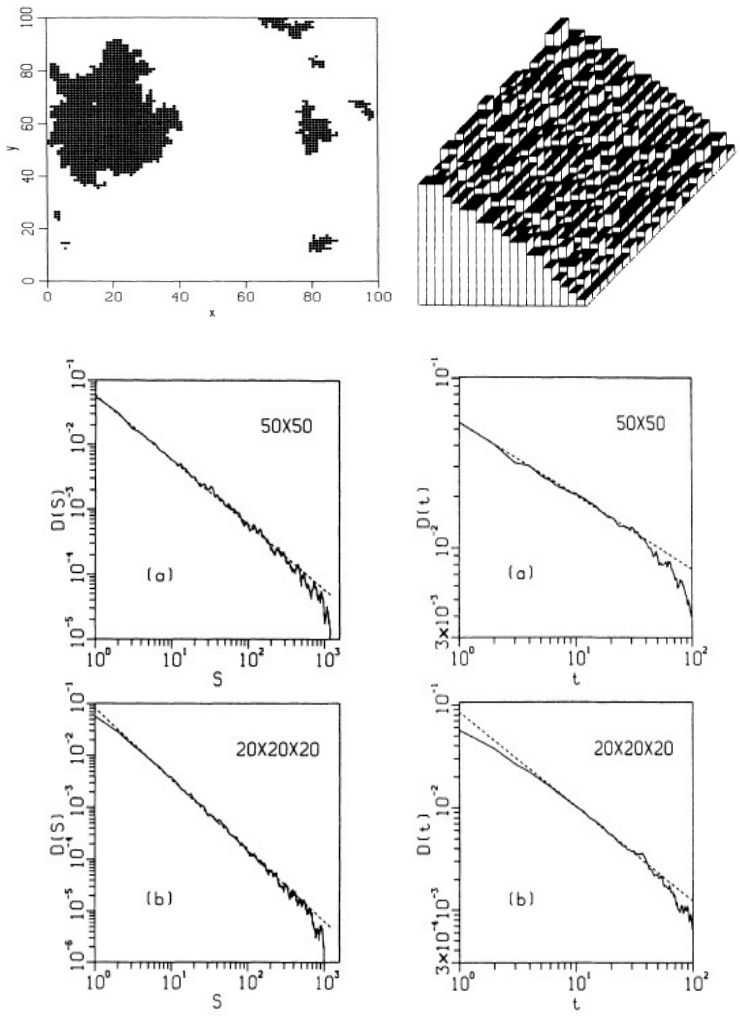


Fig. 2.2: Examples of a fragmented avalanche (top left) occurring in a 2-D (computer) sandpile (top right) and occurrence frequency distribution of avalanche cluster sizes (left panels) and avalanche durations (right panels) of the original BTW sandpile cellular automaton simulation. The simulations have been performed for a 50×50 2-D lattice (middle panels) and for a 20×20×20 3-D lattice grid (bottom panels). The powerlaw slopes are $\alpha_S = 1.0$ and $\alpha_T = 0.42$ for the 2-D grid (middle panels) and $\alpha_S = 1.37$ and $\alpha_T = 0.92$ for the 3-D grid. Reprinted from Bak, Tang, and Wiesenfeld (1987, 1988) with permission; Copyright by American Physical Society.

2.1.2 Physical Aspects

Every theoretical model should be subjected to experiments and observations. Since a cellular automaton model is a purely mathematical model of

complexity, similar to the mathematical definition of a fractal dimension, its predictions of a particular mathematical distribution function (such as a powerlaw) may represent a universal property of SOC processes, but does not represent a complete physical model *per se* that can be used for applications to real observations. The physical aspect of a SOC model is hidden in the "place-holder" variable $z_{i,j,k}$ and its scaling laws with observables, which require a specific physical mechanism for each observed phenomenon (for examples see second column in Table 2.1).

If we go back to the original SOC concept of avalanches in a sandpile, say in the 1D-version, the place-holder variable z_i has been interpreted as a vertical altitude or height difference $z_i = (h_{i+1} - h_i)/\Delta x$ between nearest-neighbor nodes, so that the instability threshold corresponds to a critical slope $z_i > z_{\rm crit} = (dh/dx)_{\rm crit}$. For the 3D version, the re-distribution rule is given in Eq. (2.1). The instability threshold essentially corresponds to a critical point where the gravitational force exceeds the frictional force of a sand grain.

In applications to solar flares or magnetospheric substorms, the "place-holder" variable $z_{i,j,k}$ has been related to the magnetic field $B_{i,j,k} = B(x_{i,j,k})$ at location $x_{i,j,k}$, and a related magnetic energy $E_B = B^2/(8\pi)$ can be defined. The instability criterion $\Delta B > z_{\rm crit}$ involves than a magnetic gradient or magnetic field curvature (Charbonneau et al. 2001), also called the *lattice Laplacian*,

$$\Delta B = B_{i,j,k} - \frac{1}{2S} \sum_{nn=1}^{2S} B_{nn} \ , \qquad |\Delta B| > z_{\rm crit} \tag{2.2}$$

where the summation includes all $2 \times S$ nearest neighbors ("nn") in a Cartesian S-dimensional lattice.

More realistically, physics-based models have been attempted by applying the magneto-hydrodynamic (MHD) equations to the lattice field $B(\mathbf{x})$, obeying Ampère's law for the current density $\mathbf{j}$,

$$\mathbf{j} = \frac{1}{4\pi}(\nabla \times \mathbf{B}) \ , \tag{2.3}$$

which yields together with Ohm's law the induction equation,

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^{\mathbf{2}} \mathbf{B} \ , \tag{2.4}$$

and fulfills the divergence-free condition for the magnetic field,

$$\nabla \cdot \mathbf{B} = 0 \ . \tag{2.5}$$

The instability threshold can then be expressed in terms of a critical resisitivity η. Such cellular models with discretized magneto-hydrodynamics (MHD) have been applied to magnetospheric phenomena, triggering magnetospheric substorms by perturbations in the solar wind (Takalo et al. 1993; 1999), as

well as to solar flares (Vassiliadis et al. 1998; Isliker et al. 1998). Initially, solar flares were modeled with isotropic magnetic field cellular automaton models (Lu and Hamilton 1991). However, since the plasma-β parameter (i.e., the ratio of the thermal to the magnetic pressure) is generally smaller than unity in the solar corona, particle and plasma flows are guided by the magnetic field, and thus SOC avalanches are more suitably represented with anisotropic 1D transport (Vlahos et al. 1995). Other incarnations of cellular SOC models mimic the magnetic field braiding that is thought to contribute to coronal heating (Morales and Charbonneau 2008). On the other hand, SOC models that involve discretized ideal MHD equations have been criticized to be inadequate to describe the highly resistive and turbulent evolution of magnetic reconnection processes, which are believed to be the driver of solar flares.

Besides the magnetic field approach of the a physical variable in SOC lattice simulations, which predicts the magnetic energy $E_B \propto B^2/(8\pi)$ or current $\mathbf{j} \propto \nabla \times \mathbf{B}$ in each node point, we still have to model the physical observables. The magnetic field or electric currents can only be observed by in-situ measurements, which is possible for magnetospheric or heliospheric phenomena with spacecraft, but often observables can only be obtained by remote-sensing measurements in form of photon fluxes emitted in various wavelength ranges, such as for solar or astrophysical sources. This requires physical modeling of the radiation process, a component that has been neglected in most previous literature on SOC phenomena.

If the physical quantity $z_{i,j,k}$ at each cellular node is defined in terms of an energy e, the discretized change of the quantity dz/dt corresponds then to a quantized amount of dissipated energy $\langle \Delta e \rangle$ per node. The instantaneous energy dissipation rate $de/dt = \langle \Delta e \rangle dV/dt$ of the system scales then with the instantaneous volume change $dV(t)/dt$, while the total amount of dissipated energy during an avalanche event can be computed from the time integral, $E = \langle \Delta e \rangle \int_0^t V(t)dt$, assuming that a mean energy quantum $\langle \Delta e \rangle$ is dissipated per unstable node.

In astrophysical applications, the observed quantity is usually a photon flux, and thus the energy quantity $e = h\nu$ corresponds to the photon energy at frequency ν, while $\langle \Delta e \rangle = Nh\nu/\Delta t$ is the photon flux of N photons that are radiated in a volume cell ΔV during a time step Δt for a given radiation process. The conversion of the intrinsic energy dissipation rate $de/dt = \langle \Delta e \rangle (dV/dt)$ (e.g., of magnetic energy E_B in a magnetic reconnection process) has then to be related to the amount of emitted photons $\langle \Delta e \rangle = Nh\nu/\Delta t$ by a physical model of the dissipation process. The physical model of the radiation process may involve a nonlinear scaling between the amount of dissipated energy and the number of emitted photons in a specific wavelength range, which changes the powerlaw slope of the observed distributions of photon fluxes, compared with the predictions of the generic quantity $z_{i,j,k}$ used in SOC cellular automaton simulations (e.g., see section 13.1.4 in chapter 13 on "SOC systems in astrophysics" in this book).

2.2 Analytical SOC Models

The complexity of the spatio-temporal pattern of a SOC avalanche is produced by a simple mathematical re-distribution rule that defines the nearest-neighbor interactions on a microscopic level, while the observed structure manifests itself at the macroscopic level. While cellular automatons operate on the microscopic level of a lattice cell, the complex macroscopic structure cannot be analytically derived from the microscopic states of the nonlinear system. In classical thermodynamics, the macroscopic state of a gas, such as the velocity distribution function of molecules can be derived from the (binomial or Gaussian) probability distribution function of microscopic states. In contrast, nonlinear dissipative systems in the state of self-organized criticality seem to exhibit a higher level of complexity, so that their macroscopic morphology cannot be analytically derived from the microscopic states. Therefore, analytical models can describe the macroscopic SOC parameters only by simplified approximations of the SOC dynamics, which are aimed to be consistent with the observed (powerlaw-like) probability or occurrence frequency distributions. We will describe a few of such analytical models in the following, which make quantitative predictions of the powerlaw slopes and scaling laws between SOC parameters. Such analytical models can then be used for Monte-Carlo simulations of SOC avalanches and be forward-fitted to the observed distributions of SOC parameters.

2.2.1 Exponential-Growth SOC Model (EG-SOC)

SOC avalanches have always a growth phase with nonlinear characteristics, similar to a multiplicative chain-reaction. Mathematically, the simplest function that grows in a multiplicative way is the exponential function with a positive growth rate. It is therefore not surprising that the first analytical models of SOC processes were based on such an exponential growth function (even when they were not called SOC models at the pre-Bak time). The attractive feature for SOC applications is the fact that the simple assumption of an exponential growth curve combined with random durations automatically leads to a powerlaw distribution function, as we will see in the following.

The earliest analytical models in terms of an exponential growth phase with saturation after a random time interval go back to Willis and Yule (1922) who applied it to geographical distributions of plants and animals. Yule's model was applied to cosmic rays (Fermi 1949), to cosmic transients and solar flares (Rosner and Vaiana 1978; Aschwanden et al. 1998), to the growth dynamics of the world-wide web (Huberman and Adamic 1999), as well as to the distribution of the sizes of incomes, cities, internet files, biological taxa, and in gene family and protein family frequencies (Reed and Hughes 2002).

Here we describe the analytical derivation of the exponential-growth SOC model, following Aschwanden (2011a, Section 3.1). We define the time evolution of the energy release rate $W(t)$ of a nonlinear process that starts at a threshold energy of W_0 by

$$W(t) = W_0 \ \exp\left(\frac{t}{\tau_G}\right) , \qquad 0 \leq t \leq \tau \ , \tag{2.6}$$

where τ_G represents the e-folding growth time of the exponential function. The process grows exponentially until it saturates at time $t = \tau$ with a saturation energy W_S,

$$W_S = W(t = \tau) = W_0 \ \exp\left(\frac{\tau}{\tau_G}\right) . \tag{2.7}$$

We define a peak energy release rate P that represents the maximum energy release rate W_S, after subtraction of the threshold energy W_0, that corresponds to the steady-state energy level before the nonlinear growth phase,

$$P = W_S - W_0 = W_0 \left[\exp\left(\frac{\tau}{\tau_G}\right) - 1\right] \ . \tag{2.8}$$

In the following, we will refer to the peak energy release rate P also briefly as "peak energy". For the saturation times τ, which we also call "rise times", we assume a random probability distribution, approximated by an exponential function $N(\tau)$ with e-folding time constant t_S,

$$N(\tau)d\tau = \frac{N_0}{t_S} \exp\left(-\frac{\tau}{t_S}\right)d\tau \ . \tag{2.9}$$

This probability distribution is normalized to the total number of N_0 events.

In order to derive the probability distribution $N(P)$ of peak energy release rates P, we have to substitute the variable of the peak energy, P, into the function of the rise time $\tau(P)$, which yields (using the functional relationship $\tau(P)$ from Eq. 2.8),

$$N(P)dP = N(\tau)d\tau = N[\tau(P)] \left|\frac{d\tau}{dP}\right| dP = \frac{N_0(\alpha_P - 1)}{W_0} \left(\frac{P}{W_0} + 1\right)^{-\alpha_P} \ dP \ , \tag{2.10}$$

which is an exact powerlaw distribution for large peak energies ($P \gg W_0$) with a powerlaw slope α_P of

$$\alpha_P = \left(1 + \frac{\tau_G}{t_S}\right) \ . \tag{2.11}$$

The powerlaw slope thus depends on the ratio of the growth time to the e-folding saturation time, which is essentially the average number of growth

times. Examples of time series with avalanches of different growth times ($\tau_G/t_S = 0.5, 1.0, 2.0$) are shown in Fig. 2.3, along with the corresponding powerlaw distributions of peak energies P. Note that the fastest growing events produce the flattest powerlaw distribution of peak energies.

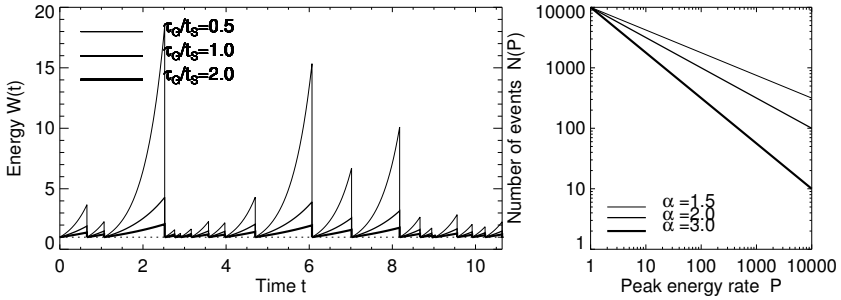


Fig. 2.3: Time evolution of energy release rate $W(t)$ for 3 different ratios of growth times to saturation times, $\tau_G/t_S = (0.5, 1.0, 2.0)$ (left) and the corresponding powerlaw distributions of the peak energy release rate P. Note that the event set with the shortest growth time ($\tau_G/t_S = 0.5$) reaches the highest energies and thus produces the flattest powerlaw slope ($\alpha = 1 + \tau_G/t_S = 1.5$).

Once an instability has released a maximum amount W_S of energy, say when an avalanche reaches its largest velocity on a sandpile, the energy release gradually slows down until the avalanche comes to rest. For sake of simplicity we assume a linear decay phase of the released energy (Fig. 2.4),

$$W(t) = W_0 + (W_S - W_0)\left(1 - \frac{(t - t_1)}{D}\right) \quad t_1 < t < t_2 \ , \tag{2.12}$$

where t_2 is the end time of the process at $t_2 = t_1 + D$. The time interval T of the total duration of the avalanche process is then the sum of the exponential rise phase τ and the linear decay phase D as illustrated in Fig. 2.4,

$$T = \tau + D = \tau_G \ln\left(\frac{P}{W_0} + 1\right) + \tau_D \frac{P}{W_0} \ . \tag{2.13}$$

We see that this relationship predicts an approximate proportionality of $T \propto P$ for large avalanches, since the second term, which is linear to P, becomes far greater than the first term with a logarithmic dependence ($\propto \ln P$).

For the calculation of the distribution $N(\tau)$ we express the total duration T in terms of the rise time τ and find a powerlaw function for the distribution of flare durations T,

$$N(T)dT = N[\tau(T)]\left|\frac{d\tau}{dT}\right| dT = \frac{N_0(\alpha_T - 1)}{\tau_D}\left(\frac{T}{\tau_D} + 1\right)^{-\alpha_T} \ dT \ . \tag{2.14}$$

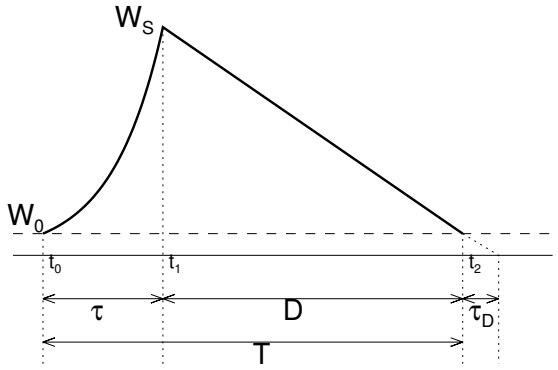


Fig. 2.4: Schematic of the time evolution of an avalanche event, consisting of (i) a rise time (τ) with exponential growth of the energy release $W(t)$ from a threshold level W_0 to the saturation level W_S, and (ii) a decay time (D) with a constant decay rate $\eta = dW/dt = W_0/\tau_D$.

We define also the total released energy E by the time integral of the energy release rate $W(t)$ during the event duration T, but neglect the rise time τ (i.e., $T \approx D$) and subtract the threshold level W_0 before the avalanche,

$$E = \int_0^T [W(t) - W_0] \; dt \approx \int_\tau^{\tau+D} [W(t) - W_0] \; dt = \frac{1}{2} PD \; . \tag{2.15}$$

leading to a powerlaw-like function for the frequency distribution of energies E,

$$N(E)dE = N[P(E)] \left| \frac{dP}{dE} \right| dE = \frac{N_0(\alpha_P - 1)}{2E_0} \left[\sqrt{\frac{E}{E_0}} + 1 \right]^{-\alpha_P} \left[\frac{E}{E_0} \right]^{-1/2} dE \tag{2.16}$$

Thus, we find the following approximate scaling laws between the powerlaw exponents,

$$\begin{aligned} \alpha_P &= 1 + \tau_G/t_S \\ \alpha_T &= \alpha_P \\ \alpha_E &= (\alpha_P + 1)/2 \end{aligned} \tag{2.17}$$

In summary, this model predicts powerlaw distribution functions for the three SOC parameters P, E, and T, which match the simulated distributions with cellular automaton simulations, as well as the observed distributions of solar flare hard X-ray fluxes (Lu and Hamilton 1991). For a ratio of $\tau_G/t_S = 1$, this model predicts $\alpha_P = 2.0$, $\alpha_T = 2.0$, and $\alpha_E = 1.5$. This particular time ratio of $\tau_G/t_S = 1$ implies that an avalanche typically saturates after one exponential growth time (see the spatio-temporal patterns for small avalanches

in Fig. 2.5). Open questions are: Which physical process would explain this particular time scale ratio $\tau_G/t_S = 1$? What is the physical interpretation of the linear decay phase? Is the exponential distribution of avalanche growth times consistent with observations, since powerlaw-like distributions are expected for SOC parameters? How can the observed intermittent fluctuations of time profiles and the fractal geometry be accomodated in a model with a monotonic growth function? Although this model seems not to reproduce all observed properties of SOC phenomena, it has a didactical value, since it represents the most basic model that links the nonlinear evolution of instabilities to the powerlaw distributions observed in SOC phenomena.

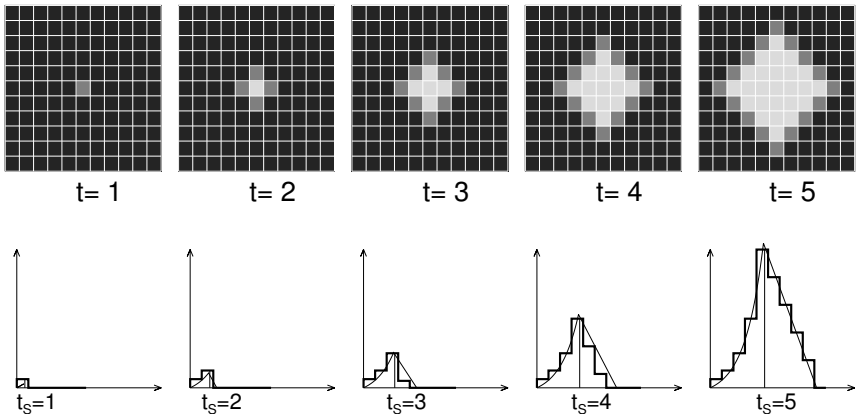


Fig. 2.5: Spatial patterns of a propagating avalanche in subsequent time steps in a 2-D cellular automaton model with a nearest-neighbor redistribution rule (top) and time profiles of energy release rate (bottom), for saturation times of $t_S = 1, 2, ..., 5\Delta t$. The black cells represent cells with random fluctuations below the threshold, $z_k < z_c$, the grey cells contain possibly unstable cells with fluctuation $z_k \geq z_c$ that are subject to a first redistribution, while the white cells have already been affected by a redistribution rule before. Most avalanches die out after step t $\gtrsim 2$.

The exponential-growth model is most suitable for multiplicative avalanche processes, where the increase per time step during the rise phase is based on a multiplicative factor, such as it occurs in nuclear chain reactions, population growth, or urban growth. Alternatively, avalanche processes that continuously expand in space may show an energy increase that scales with the area or volume, which has a powerlaw relationship to the spatial or temporal scale, i.e., $A(t) \propto r(t)^2 \propto t^2$ or $V(t) \propto r(t)^3 \propto t^3$. Such a model with a powerlaw-growth function

$$W(t) = W_0 \left[1 + \left(\frac{t}{\tau_G}\right)^p\right] , \tag{2.18}$$

rather than an exponential-growth function (Eq. 2.6) was computed in Aschwanden (2011a; Section 3.2), which predicts identical scaling laws between the SOC parameters (P, E, T), but the occurrence frequency distributions exhibit an exponential fall-off at the upper end. Otherwise it has similar caveats as the exponential-growth model (i.e., monotonic growth curve rather than intermittency, and Euclidean rather than fractal avalanche volume).

Another variant of the exponential-growth model is the logistic-growth model (Aschwanden 2011a; Section 3.3), which has a smoother transition from the initially exponential growth to the saturation phase, following the so-called *logistic equation* (discovered by Pierre François Verhulst in 1845; see review by Yule 1925 or textbooks on nonlinear dynamics, e.g., May 1974; Beltrami 1987, p.61; and Jackson 1989, p.75),

$$\frac{dE(t)}{dt} = \frac{E(t)}{\tau_G} \cdot \left[1 - \frac{E(t)}{E_\infty}\right] \ , \tag{2.19}$$

where the dissipated energy is limited by the so-called *carrying capacity* limit E_∞ used in ecological applications. The resulting occurrence frequency distributions are powerlaws for the energy E and peak energy dissipation rate P, but exponential functions for the rise time τ and duration T of an avalanche. Like the exponential-growth and the powerlaw-growth model, the temporal intermittency and the geometric fractality observed in real SOC phenomena are not reproduced by the smoothly-varying time evolution of the logistic-growth model.

2.2.2 The Fractal-Diffusive SOC Model (FD-SOC)

The microscopic structure of a SOC avalanche has been simulated with a discretized mathematical re-distribution rule, which leads to highly inhomogeneous, filamentary, and fragmented topologies during the evolution of an avalanche. Therefore, it appears to be adequate to develop an analytical SOC model that approximates the inhomogeneous topology of an avalanche with a fractal geometry. Bak and Chen (1989) wrote a paper entitled "The physics of fractals", which is summarized in their abstract with a single sentence: *Fractals in nature originate from self-organized critical dynamical processes.* Thus, we introduce now an analytical model that includes the fractal geometry and a diffusion-type spatio-temporal transport process, which we call the fractal-diffusive SOC model (FD-SOC).

Such a statistical fractal-diffusive avalanche model of a slowly-driven SOC system has been derived in Aschwanden (2012). This analytical model represents a universal (physics-free) description of the statistical time evolution and occurrence frequency distribution function of SOC processes. It is based on four fundamental assumptions: (1) A SOC avalanche grows spatially like

a diffusive process; (2) The spatial volume of the instantaneous energy dissipation rate is fractal; (3) The time-averaged fractal dimension is the mean of the minimum dimension $D_{S,min} \approx 1$ (for a sparse SOC avalanche) and the maximum dimension $D_{S,max} = S$ (given by the Euclidean space); and (4) The occurrence frequency distribution of length scales is reciprocal to the size L of spatial scales, i.e., $N(L) \propto L^{-S}$ in Euclidean space with dimension S (also called the *scale-free probability conjecture*). We will discuss these assumptions in more detail in the following.

The first assumption of a diffusive process is based on numerical simulations of cellular automaton models. A SOC avalanche propagates in a cellular automaton model by nearest-neighbor interactions in a critical state, where energy dissipation propagates only to the nearest-neighbor cells (in a S-dimensional lattice grid) that are above a critical threshold. This mathematical rule that describes the entire dynamics and evolution of a SOC avalanche is very simple for a single time step, but leads to extremely complex spatial patterns after a finite number of time steps. For a visualization of a large number of such complex spatial patterns generated by a simple iterative mathematical redistribution rule see, for instance, the book "A New Kind of Science" by Wolfram (2002). The complexity of these spatial patterns can fortunately be approximated with a single number, the fractal dimension D_S. If one monitors the time evolution of a spatial pattern of a SOC avalanche in a cellular automaton model, one finds that the length scale $x(t)$ evolves with time approximately with a diffusive scaling (see radius $r(t)$ of snapshots of a 2-D cellular automaton evolution in Fig. (2.6) and its time evolution $r(t) \propto t^{1/2}$ in Fig. (2.8), bottom right panel),

$$x(t) \propto t^{1/2} \ , \tag{2.20}$$

which leads to a statistical scaling law between the avalanche sizes $L = x(t = T)$ and time durations T of SOC avalanches,

$$L \propto T^{1/2} \ . \tag{2.21}$$

The second assumption of a fractal pattern of the instantaneous energy rate is also based on tests with cellular automaton simulations (see measured fractal dimensions D_2 of snapshots in Fig. (2.7) and the time evolution $D_2(t)$ in Fig. (2.8) top right panel). The fractal dimension is essentially a simplified parameter that describes the "micro-roughness", "graininess", or inhomogeneity of critical nodes in a lattice grid in the state of self-organized criticality. Of course, such a single number is a gross over-simplification of a complex system with a large number of degrees of freedom, but the numerical simulations confirm that avalanche patterns are fractal (Fig. 2.7). Thus we define the volume $V_S(t)$ of the instantaneous energy dissipation rate (i.e., the number of "active sites" or nodes) in terms of a fractal (Hausdorff) dimension D_S that scales with the length scale x as,

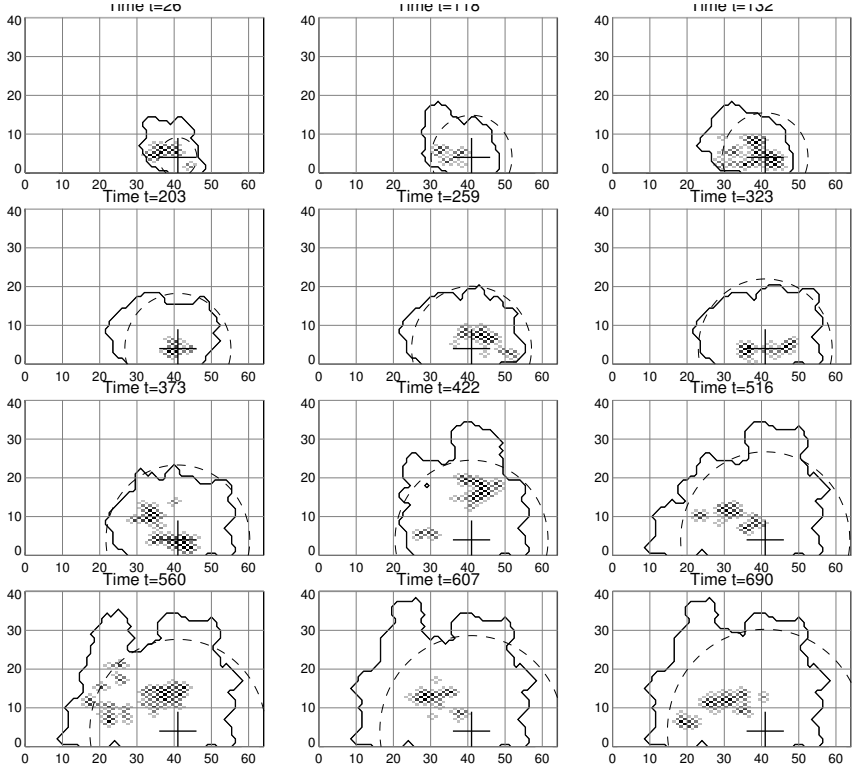


Fig. 2.6: Time evolution of a large avalanche event in a 2-D cellular automaton simulation with grid size $N = 64^2$. The 12 panels show snapshots taken at a subset of peak times, from $t = 26$ to $t = 690$, when the energy dissipation rate reached a local maximum as a function of time. Active nodes where energy dissipation occurs at time t are visualized with black and grey points, depending on the energy dissipation level. The starting point of the avalanche occurred at pixel $(x, y) = (41, 4)$, which is marked with a cross. The time-integrated envelop of the avalanche is indicated with a solid contour, and the diffusive avalanche radius $r(t) = t^{1/2}$ is indicated with a dashed circle (Aschwanden 2012).

$$V_S(t) \propto x^{D_S} \ , \tag{2.22}$$

which leads also to a statistical scaling law between avalanche volumes V and spatial scales L or durations T of SOC avalanches (with Eq. 2.21)

$$V_S \propto L^{D_S} \propto T^{D_S/2} \ . \tag{2.23}$$

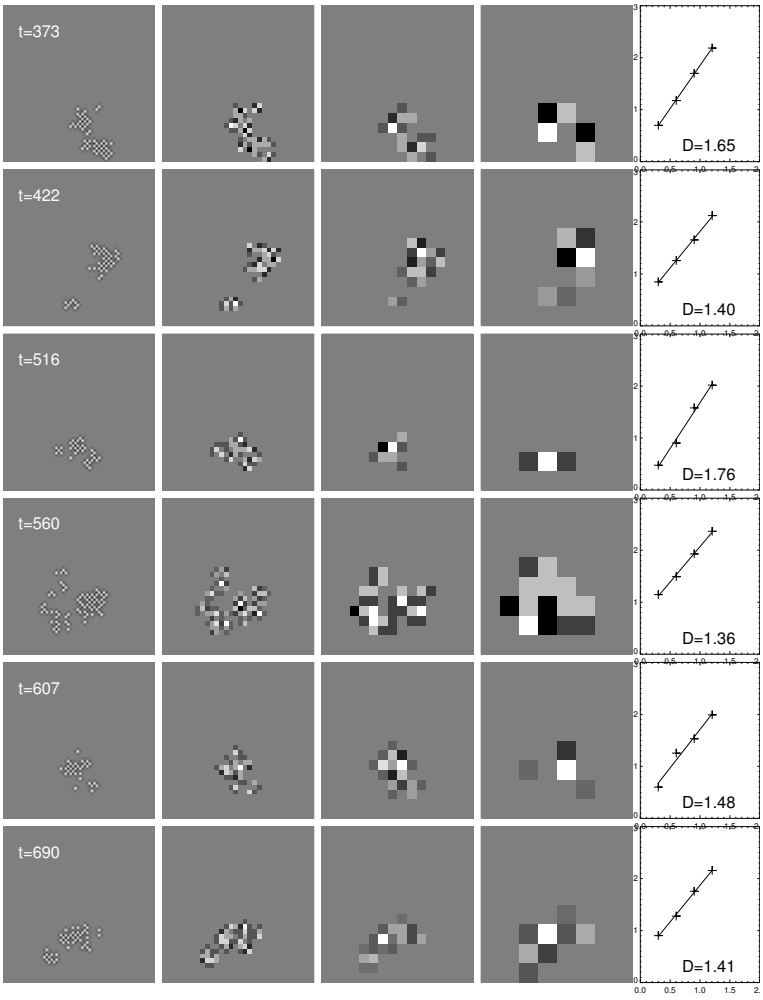


Fig. 2.7: Determination of the fractal dimension $D_2 = \log A_i / \log x_i$ for the instantaneous avalanche sizes of 6 time steps of the avalanche event shown in Fig. (2.6). Each row is a different time step and each column represents a different binning of macropixels ($\Delta x_i = 1, 2, 4, 8$). The fractal dimension is determined by a linear regression fit shown on the right-hand side. The mean fractal dimension of the 12 avalanche snapshots shown in Fig. (2.6) is $D_2 = 1.43 \pm 0.17$ (Aschwanden 2012).

The third assumption of the mean fractal dimension has also been confirmed by numerical simulations of cellular automaton SOC processes in all three dimensions $S = 1, 2, 3$ (Aschwanden 2012), but it can also be under-

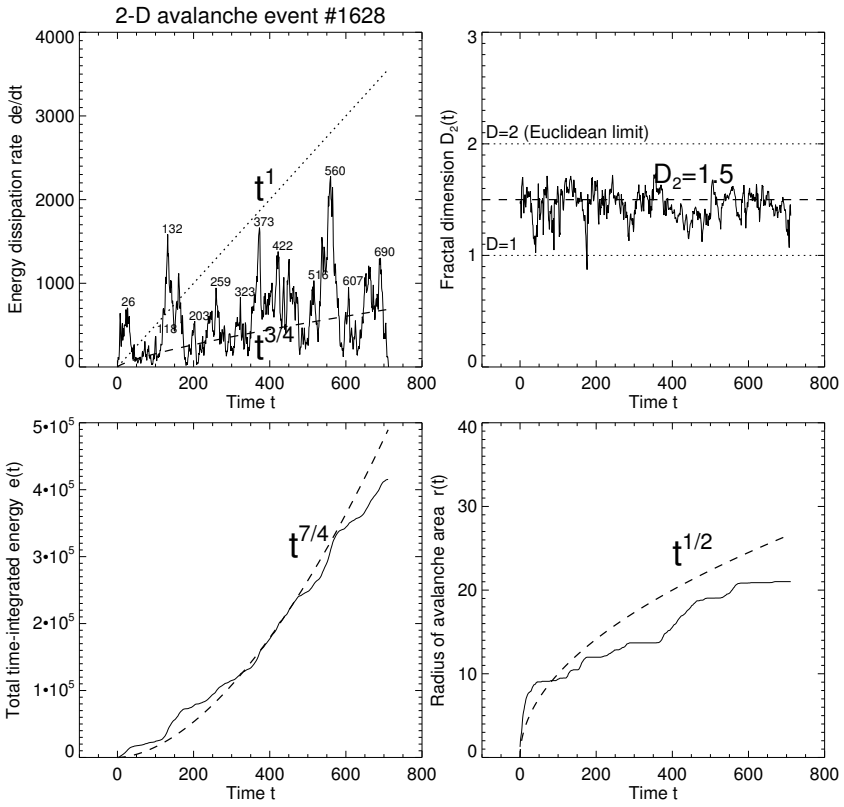


Fig. 2.8: Time evolution of the same large avalanche event from a 2-D cellular automaton simulation with grid size $N = 64^2$ as shown in Fig. (2.6). The time profiles include the instantaneous energy dissipation rate $f(t) = de/dt$ (top left), the time-integrated total energy $e(t)$ (bottom left), the instantaneous fractal dimension $D_2(t)$ (top right), and the radius of the avalanche area $r(t)$ (bottom right). The observed time profiles from the simulations are outlined in solid linestyle and the theoretically predicted average evolution in dashed linestyle. The statistically predicted values of the instantaneous energy dissipation rate $f(t) \propto t^{3/4}$ (dotted curve) and peak energy dissipation rate $p(t) \propto t^1$ (dashed curve) after a time interval t are also shown (top left panel). The 12 time labels from 26 to 690 (top left frame) correspond to the snapshot times shown in Fig. (2.6) (Aschwanden 2012).

stood by the following plausibility argument. The sparsest SOC avalanche that propagates by nearest-neighbor interactions is the one that spreads only in one spatial dimension, and thus yields an estimate of the minimum fractal dimension of $D_{S,min} \approx 1$, while the largest SOC avalanche is almost

space-filling and has a volume that scales with the Euclidean dimension, $D_{S,max} = S$. Combining these two extremal values, we can estimate a time-averaged fractal dimension $\langle D_S \rangle$ from the arithmetic mean,

$$\langle D_S \rangle \approx \frac{D_{S,min} + D_{S,max}}{2} = \frac{1+S}{2} \ , \tag{2.24}$$

which yields a mean fractal dimension of $\langle D_3 \rangle = (1+3)/2 = 2.0$ for the 3D case ($S = 3$).

The fourth assumption on the size distribution of length scales is a probability argument. The system size L_{sys} of a SOC system represents an upper limit of spatial scales L for SOC avalanches, i.e., $L \leq L_{sys}$. For the 3D-case, the volumes V of individual avalanches are bounded by the volume V_{sys} of the system size, i.e., $V = L^3 \leq V_{sys} = L_{sys}^3$. If the entire system is in a critical state, the maximum number of SOC avalanches that can be produced throughout the system is given by the packing density, and the probability $N(L)$ for a fixed avalanche size L with volume V is simply reciprocal to the size, (visualized in Fig. 2.9), i.e.,

$$N(L) \propto \frac{V_{sys}}{V(L)} \propto V(L)^{-1} \ , \tag{2.25}$$

which is equivalent to

$$N(L)dL \propto \frac{L_{sys}^3}{L^3} dL \propto L^{-3} dL \ . \tag{2.26}$$

In reality, only one single or no avalanche occurs at a given time, but the relative probability for a small or large avalanches size still scales with the reciprocal size, if we use a probabilistic argument (as it is used in the derivation of a binomial distribution by adding up all possible combinations and permutations that produce a particular outcome).

Based on these four model assumptions we can now quantify the time evolution of SOC parameters. For sake of simplicity we apply this fractal-diffusive SOC model to an astrophysical source that emits a photon flux $f(t)$ that is proportional to the instantaneous energy dissipation volume $V_S(t)$ of a SOC avalanche, where a mean energy quantum $\langle \Delta E \rangle$ is emitted per volume cell element ΔV. Thus, the emitted flux is (using the diffusive scaling $x(t) \propto t^{1/2}$),

$$f(t) = \frac{de(t)}{dt} \propto \langle \Delta E \rangle V_S(t) = \langle \Delta E \rangle x(t)^{D_S} = \langle \Delta E \rangle t^{D_S/2} \ . \tag{2.27}$$

In the 2-D case with $D_S = D_2 = (1+2)/2 = 3/2$ (Eq. 2.24) we expect a statistical scaling of $f(t) \propto t^{3/4}$ (see Fig. 2.8 top left). In the 3-D case with $D_S = D_3 = (1+3)/2 = 2$, we expect than the proportionality $f(t) \propto t^{D_3/2} \propto t^1$. In Fig. (2.10, bottom panel) we simulate such a flux time profile $f(t)$

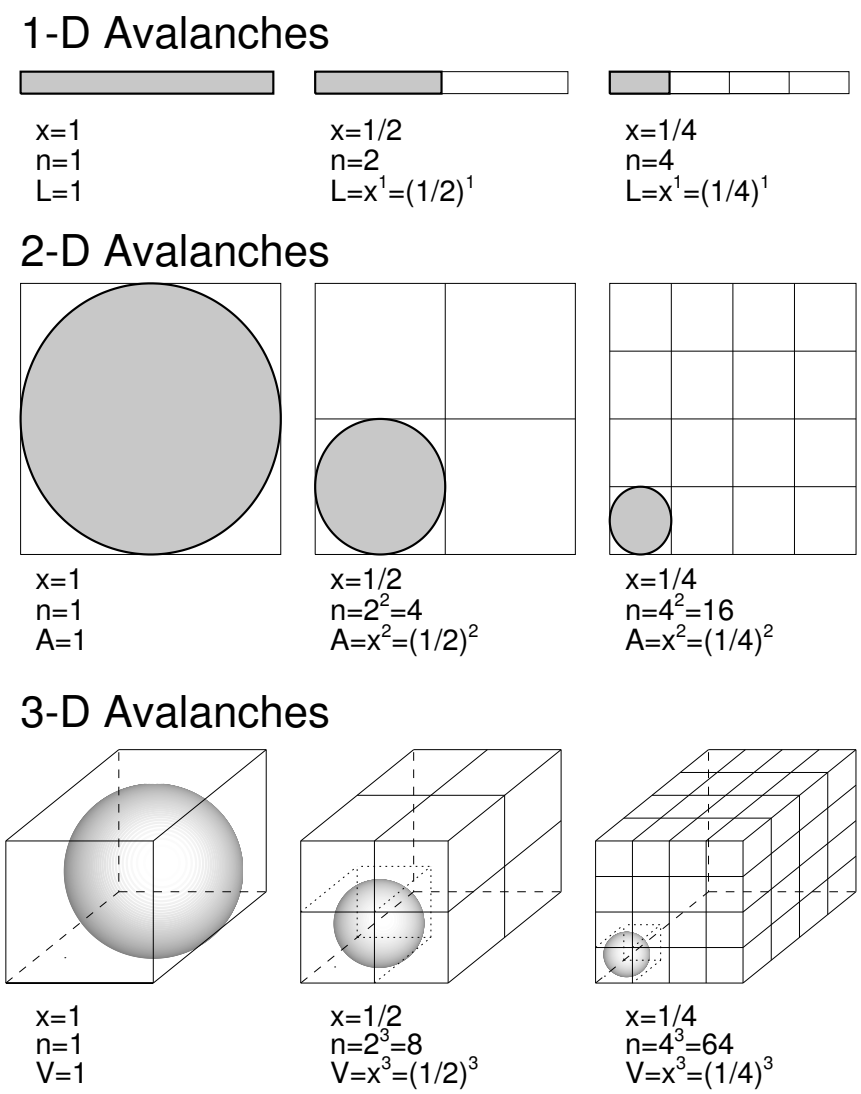


Fig. 2.9: Schematic diagram of the Euclidean volume scaling of the diffusive avalanche boundaries, visualized as circles or spheres in the three Euclidean space dimensions $S = 1, 2, 3$. The Euclidean length scale x of subcubes decreases by a factor 2 in each step ($x_i = 2^{-i}, i = 0, 1, 2$), while the number of subcubes increases by $n_i = (2^i)^S$, defining a probability of $N(x_i) \propto x_i^{-S}$ for each avalanche size with size x_i.

by applying noise fluctuations in the time evolution of the fractal dimension $D_2(t)$ (Fig. 2.10, top).

The statistical peak value $p(t)$ of the energy dissipation rate after time t can be estimated from the largest possible avalanches, which have an almost space-filling dimension $D_S \lesssim S$,

$$p(t) = \frac{de(t)}{dt} \propto \langle \Delta E \rangle V_{S,max}(t) = \langle \Delta E \rangle x(t)^S = \langle \Delta E \rangle t^{S/2} \ . \tag{2.28}$$

and thus would be expected to scale as $p(t) \propto t^{S/2} \propto t^{1.5}$ for the $S = 3$ case. Thus the time profile of peak values envelopes the maximum fluctuations of the flux time profile $f(t)$ (Fig. 2.10, second panel).

The evolution of the total dissipated energy $e(t)$ after time t is simply the time integral, for which we expect

$$e(t) = \int_0^t \frac{de(\tau)}{d\tau} d\tau \propto \int_0^t \tau^{D_S/2} \propto t^{(1+D_S/2)} \ , \tag{2.29}$$

which yields the function $e(t) \propto t^{7/4}$ for the 2-D case (Fig. 2.8, bottom left), and a function $e(t) \propto t^2$ for the 3-D case (Fig. 2.10, third panel). These time evolutions apply to every SOC model that has an emission $f(t)$ proportional to the fractal avalanche volume $V_S(t)$. For applications to observations in a particular wavelength range there may be an additional scaling law between the avalanche volume and emission (or intensity) of the underlying radiation mechanism.

In order to derive the size distributions of these various observables, we start with the probability distribution of (avalanche) length scales (Eq. 2.25),

$$N(L) \propto V_S^{-1} \propto L^{-S} \ . \tag{2.30}$$

which follows from a simple statistical probability argument as discussed in the derivation of Eq. (2.26). This is an extremely important assumption, which automatically predicts a powerlaw distribution for length scales. Using the scaling laws that result from Eq. (2.21) and (2.27)-(2.29) for characteristics scales $L = x(t = T)$, $E = e(t = T)$, $F = f(t = T)$, and $P = p(t = T)$ for an avalanche duration time T,

$$\begin{array}{l} L \propto T^{1/2} \\ F \propto T^{D_S/2} \\ P \propto T^{S/2} \\ E \propto T^{1+D_S/2} \end{array} \ , \tag{2.31}$$

we can directly calculate the occurrence frequency distributions for all these SOC parameters, by substituting the variables from the correlative relationships given in Eq. (2.31), yielding

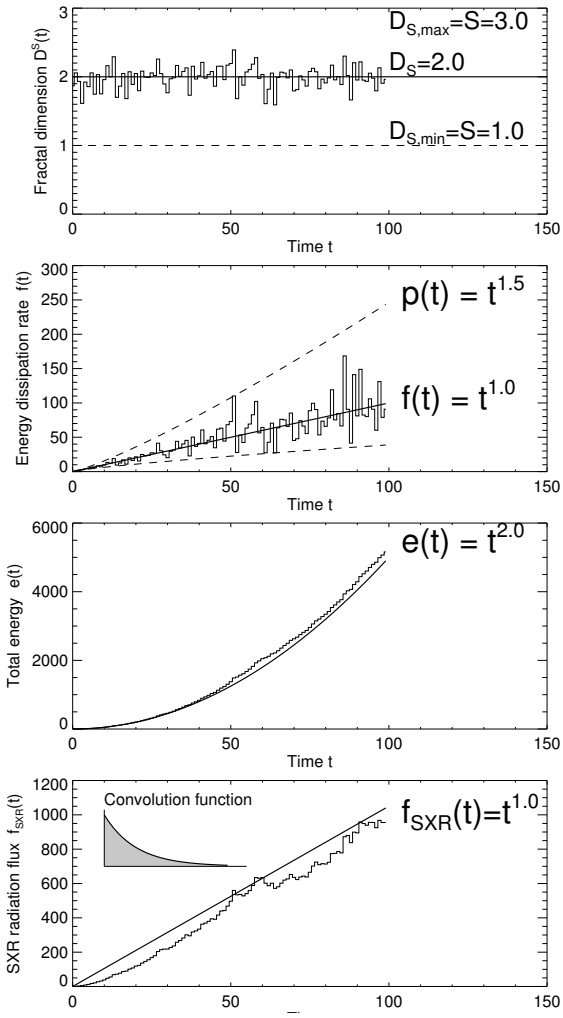


Fig. 2.10: Simulation of the fractal-diffusive SOC model for an Euclidean dimension $S = 3$, showing the time evolution of the fractal dimension $D_S(t)$ (top panel), the instantaneous energy dissipation rate $f(t)$ and peak energy dissipation rate $p(t)$ (second panel), the total time-integrated dissipated energy $e(t)$ (third panel), and the soft X-ray time profile $f_{\rm sxr}(t)$ (bottom panel), which results from the convolution of the instantaneous energy dissipation rate $f(t)$ (second panel) with an exponential decay function with an e-folding time of $\tau_{\rm decay}$ (shown in insert of bottom panel) (Aschwanden and Freeland 2012).

$$N(T)dT = N(L[T]) \left| \frac{dL}{dT} \right| dT \propto T^{-[(1+S)/2]} \; dT \; . \tag{2.32}$$

$$N(F)dF = N(T[F]) \left| \frac{dT}{dF} \right| dF \propto F^{-[1+(S-1)/D_S]} \; dF \; , \tag{2.33}$$

$$N(P)dP = N(T[P]) \left| \frac{dT}{dP} \right| dP \propto P^{-[2-1/S]} \; dP \; , \tag{2.34}$$

$$N(E)dE = N(T[E]) \left| \frac{dT}{dE} \right| dE \propto E^{-[1+(S-1)/(D_S+2)]} \; dE \; . \tag{2.35}$$

If such simple single powerlaw scaling laws exist, this derivation yields naturally powerlaw functions for all parameters L, T, F, P, and E, which are the hallmarks of SOC systems. In summary, if we denote the occurrence frequency distribution $N(x)$ of a parameter x with a powerlaw distribution with powerlaw exponent α_x,

$$N(x)dx \propto x^{-\alpha_x} \; dx \; , \tag{2.36}$$

we have the following powerlaw exponents α_x for the parameters $x = T, F, P$, and E,

$$\begin{array}{l} \alpha_T = (1+S)/2 \\ \alpha_F = 1 + (S-1)/D_S \\ \alpha_P = 2 - 1/S \\ \alpha_E = 1 + (S-1)/(D_S+2) \end{array} \; . \tag{2.37}$$

The powerlaw exponents α_x and correlation are summarized in Table 2.2 separately for each Euclidean dimension $S = 1, 2, 3$.

These correlation coefficients and powerlaw exponents of frequency distributions have been found to agree within $\approx 10\%$ with numerical simulations of cellular automatons for all three Euclidean dimensions ($S = 1, 2, 3$) (Aschwanden 2011a). Some deviations, especially fall-offs at the upper end of powerlaw distributions, are likely to be caused by finite-size effects of the lattice grid.

2.2.3 Astrophysical Scaling Laws

SOC theory applied to astrophysical observations covers many different wavelength regimes, for instance gamma-rays, hard X-rays, soft X-rays, and extreme ultra-violet (EUV) in the case of solar flares. A comprehensive review of such studies is given in Section 7 of Aschwanden (2011a). However, since each wavelength range represents a different physical radiation mechanism, we have to combine now the physics of the observables with the (physics-free) SOC statistics.

Table 2.2: Theoretically predicted occurrence frequency distribution power-law slopes α and parameter correlations predicted for SOC cellular automatons with Euclidean space dimensions $S = 1, 2, 3$, for the fractal dimension D_S, length scale L, time duration T, instantaneous energy dissipation rate (or flux) F, peak energy dissipation rate (or peak flux) P, and total time-integrated energy E.

Theory	S=1	S=2	S=3
$D_S = (1+S)/2$	1	3/2	2
$\alpha_L = S$	1	2	3
$\alpha_T = (1+S)/2$	1	3/2	2
$\alpha_F = 1 + (S-1)/D_S$	1	5/3	2
$\alpha_P = 2 - 1/S$	1	3/2	5/3
$\alpha_E = 1 + (S-1)/(D_S+2)$	1	9/7	3/2
$L \propto T^{1/2}$	$L \propto T^{1/2}$	$L \propto T^{1/2}$	$L \propto T^{1/2}$
$F \propto T^{D_S/2}$	$F \propto T^{1/2}$	$F \propto T^{3/4}$	$F \propto T^{1}$
$P \propto T^{S/2}$	$P \propto T^{1/2}$	$P \propto T^{1}$	$P \propto T^{3/2}$
$E \propto T^{1+D_S/2}$	$E \propto T^{3/2}$	$E \propto T^{7/4}$	$E \propto T^{2}$

Let us consider soft and hard X-ray emission in solar or stellar flares. Soft X-ray emission during solar flares is generally believed to result from thermal free-free and free-bound radiation of plasma that is heated in the chromosphere by precipitation of non-thermal electrons and ions, and which subsequently flows up into coronal flare (or post-flare) loops, a process called "chromospheric evaporation process" (for a review see, e.g., Aschwanden 2004). Therefore, we can consider the flare-driven chromospheric heating rate as the instantaneous energy dissipation process of a SOC avalanche, as shown in the simulated function $f(t)$ in Fig. 2.10 (second panel). The heated plasma, while it fills the coronal flare loops, loses energy by thermal conduction and by radiation of soft X-ray and EUV photons, which generally can be characterized by an exponential decay function after an impulsive heating spike. In Fig. 2.10 (bottom) we mimic such a soft X-ray radiation light curve by convolving the instantaneous energy dissipation rate $f(t)$ (Fig. 2.10, second panel) with an exponentially decaying radiation function (with an e-folding time constant of $\tau_{\rm decay}$),

$$f_{sxr}(t) = \int_{-\infty}^{t} f(t) \exp\left[-\frac{(t-t')}{\tau_{\rm decay}}\right] dt' \ , \tag{2.38}$$

which shows also a time dependence that follows approximately

$$f_{\rm sxr}(t) \propto f(t) \propto t^{1.0} \ , \tag{2.39}$$

because the convolution with an exponential function with a constant e-folding time acts like a constant multiplier. In the limit of infinitely long decay times ($\tau_{\rm decay} \mapsto \infty$), our convolution function (Eq. 2.38) turns into a time integral of the heating function $f(t)$, which is known as Neupert effect (Dennis and Zarro 1993; Dennis et al. 2003).

The heating function is identified with the non-thermal hard X-ray emission, which indeed exhibits a highly fluctuating and intermittent time profile for energies of $\gtrsim 25$ keV, where non-thermal emission dominates (Fig. 2.8),

$$f_{\rm hxr}(t) \propto p(t) \propto t^{1.5} \ . \tag{2.40}$$

Thus the occurrence frequency distributions of fluxes are expected to be different for soft X-rays and hard X-rays. The hard X-ray flux $f_{\rm hxr}(t)$ follows the statistics of the highly fluctuating peak energy dissipation rate $p(t)$, while the soft X-ray flux $f_{\rm sxr}(t)$ is expected to follow the statistics of the smoothly-varying (convolved) time profile $f(t)$. The total duration T of energy release of an avalanche (or flare here) corresponds essentially to the rise time $t_{\rm rise}$ of the soft X-ray flux, or to the total flare duration for hard X-rays, because the decay phase of a soft X-ray flare light curve is dominated by conductive and radiative loss, rather than by continued heating input. Thus based on the generic relationships summarized in Table 2.2 we expect for the 3-D case ($S = 3$),

$$N(T) \propto T^{-\alpha_T} = T^{-2} \ , \tag{2.41}$$

$$N(f_{\rm sxr}) \propto F^{-\alpha_F} = F^{-[1+(S-1)/D_S]} = F^{-2} \ , \tag{2.42}$$

$$N(f_{\rm hxr}) \propto P^{-\alpha_P} = P^{-[2-1/S]} = P^{-5/3} \ . \tag{2.43}$$

Applications to observations did show satisfactory agreement with these theoretical values, i.e. powerlaw slopes of $\alpha_F = 2.0$ for soft X-rays and $\alpha_P = 1.67$ for hard X-rays (Aschwanden 2011a; 2011b; Aschwanden and Freeland 2012), except for the occurrence frequency distributions of flare durations T during solar cycle maxima, when the flare pile-up bias appears to have a steepening side effect (Aschwanden 2012).

2.2.4 Earthquake Scaling Laws

While the foregoing discussion is relevant to astrophysical SOC phenomena (solar and stellar flares), similar physical scaling relationships between the observables and SOC cellular automaton quantities $z_{i,j,k}$ can be developed in other fields. For earthquakes in geophysics for instance, measured quantities include the length L_S and width L_w of a rupture area, so the rupture area has the scaling $A \propto L_S L_w$. For large ruptures, however, when the surface rupture length is much larger than the rupture width, i.e., $L_s \gg L_w$, the width L_w was found to be approximately constant, so that large ruptures saturate the

faults surface and scale with a fractal dimension of $D = 1$, whereas a smaller rupture propagates across both dimension ($D = 2$) of the fault face (Yoder et al. 2012). The modeling of size distributions of earthquake magnitudes can then be conducted with a similar methodology as outlined with our fractal-diffusive model (Section 2.2.2), requiring: (i) The definition of the Euclidean space dimension for earthquakes, which could be $S = 2$ if they are treated as surface phenomena without significant depth variability, or by $S = 3$ otherwise; (ii) The measurement of the fractal dimension D, which can exhibit multi-fractal scaling from $D = 1$ for large earthquakes to $D = 2$ for small earthquakes; and (iii) assigning the observable m, which is the magnitude of the earthquake, to the corresponding physical quantity, i.e., the peak energy dissipation rate P, the smoothed energy dissipation rate F, or total time-integrated energy E. Furthermore, earthquake size distributions are often reported in terms of a cumulative distribution function $N(> m)$, which has a powerlaw index $\beta \approx \alpha - 1$ that is flatter by one than the differential distribution function (or probability density function) $N(m)\ dm$.

2.3 Alternative Models Related to SOC

Here we discuss a number of alternative dynamical models that are related to SOC models or have similar scaling laws, and discuss what they have in common with SOC or where they differ. A matrix between processes and observables is synthesized in Fig. 2.15.

2.3.1 Self-Organization Without Criticality (SO)

Self-organization (SO) often refers to geometric patterns that originate in mutual interaction of their elements, without coordination from outside the system (Camazine et al. 2001). For instance, convection arranges itself into a regular pattern of almost equal-sized convection cells, a structuring process also known as *Rayleigh-Bénard cells* (Getling 1998). Other examples are the regular wavy pattern of sand dunes in the desert (Bagnold 1941, 2005), wavy patterns of Cirrus clouds in the Earth atmosphere (Nagel and Raschke 1992), Jupiter's atmosphere with white bands of ammonia ice clouds (Antipov et al. 1985), the spiral-like patterns of the Belousov-Zhabotinsky reaction-diffusion system (Rovinsky and Menzinger 1993), or geometric patterns in biology (Camazine et al. 2001), such as the skin of zebras, giraffes, tigers, tropical fishes, or formation flight of birds. From these examples we see that self-organization mostly refers to the self-assembly of geometric patterns, which are more or less stable over long time intervals, although they arise from the physics of non-equilibrium processes, which can involve diffusion, turbulence, convec-

tion, or magneto-convection, which are governed by long-range interactions (via pressure and forces).

What is the difference to self-organized criticality (SOC)? Are sandpile avalanches a self-organizing (SO) pattern? In the standard scenario of the sandpile SOC model, individual avalanches are a local phenomenon that are randomly triggered in space and time, but occur independently, at least in the slowly-driven case. Thus, one avalanche has no mutual interaction with another avalanche and the outcome of the final size is independent of another, in contrast to self-organization without criticality, where the interaction between system-wide structures is coupled. In other words, sandpile avalanches are governed by localized disturbances via nearest-neighbor interactions, while self-organizing patterns may be formed by both nearest-neighbor and long-range interactions. Another difference is that SO creates spatial patterns, while SOC generates dynamical events in space and time (i.e., avalanches). Also the statistical distributions of the two processes are different. A self-organizing pattern is likely to produce a preferred size scale (e.g., the solar granulation or the width of zebra stripes), while self-organized criticality produces a scale-free powerlaw distribution of avalanche sizes. The difference between SO and SOC may be best illustrated with a sandpile analogy. The same sandpile can be subject to self-organization (SO), for instance when a steady wind blows over the surface and produces wavy ripples with a regular spacing pattern, as well as be subject to self-organized criticality (SOC), when intermittent avalanches occur due to random-like disturbances by infalling sand grains. The former geometric pattern may appear as a spatial-periodic pattern, while the latter may exhibit a fractal geometry.

The concept of self-organization could be also related to the intensively studied CMLs (Coupled Map Lattices), which produce chaotic maps (e.g., the Bernoulli map) on the regularly distributed nodes. The next level of complexity is captured in complex networks (complex topology with chaotic systems on the nodes). Both concepts lead also to highly complex spatio-temporal dynamics.

2.3.2 Forced Self-Organized Criticality (FSOC)

In Bak's original SOC model, which is most genuinely reproduced by sandpile avalanches and cellular automaton simulations, criticality (or the susceptible state) is continuously restored by nearest-neighbor interactions. If the local slope becomes too steep, an avalanche will erode it back to the critical value. If the erosion by an avalanche flattened the slope too much, it will be gradually restored by random input of dropped sand grains in the slowly-driven limit. When the observations were extended to magnetospheric substorms in the night-side geotail, powerlaw-like size distributions were found, suggesting a self-organized critical process, but at the same time correlations with

magnetic reconnection events at the dayside of the magnetosphere were identified, driven by solar wind fluctuations, suggesting some large-scale transport processes and long-range coupling between the day-side and night-side of the magnetosphere. Thus, these long-range interactions that trigger local instabilities in the current sheet in the Earth's magnetotail were interpreted as external forcing or loading, and a combined *forced and/or self-organized criticality (FSOC)* process was suggested (Chang 1992, 1998a,b, 1999a,b). A forced SOC process may apply to a variety of SOC phenomena in the magnetosphere, such as magnetotail current disruptions (Lui et al. 2000), substorm current disruptions (Consolini and Lui 1999), bursty bulk flow events (Angelopoulos et al. (1996, 1999), magnetotail magnetic field fluctuations (Hoshino et al. 1994), auroral UV blobs (Lui et al. 2000; Uritsky et al. 2002), auroral optical blobs (Kozelov et al. 2004), auroral electron (AE) jets (Takalo et al. 1993; Consolini 1997), or outer radiation belt electron events (Crosby et al. 2005).

What distinguishes the FSOC model from the standard (BTW-type) SOC model is mostly the long-range coupling in triggering of instabilities, which is more localized in sandpiles and cellular automaton models. The driving force in the FSOC model is a long-distance action (loading process), while it is a random local disturbance in the BTW model. Otherwise, the FSOC model entails also powerlaw distributions for the avalanche events, fractality, intermittency, statistical independence of events (indicated by random waiting time distributions), and a critical threshold (for the onset of local plasma instabilities and/or magnetic reconnection processes).

2.3.3 Brownian Motion and Classical Diffusion

Brownian motion is a concept from classical physics that describes the random motion of atoms or molecules in a gas or liquid (named after the Scottish botanist Robert Brown), which can be best observed when a neutral gas of a different color is released in the atmosphere. What we observe then is an isotropic diffusion process (in the absence of forces or flows) that monotonically increases with the square-root of time,

$$\langle x(t) \rangle \propto t^{1/2} \ . \tag{2.44}$$

This statistical trend was derived in classical thermodynamics, assuming a Gaussian distribution of velocities for the gas molecules, so that the kinetic energy of particles follow a Boltzmann distribution near thermodynamic equilibrium. The classical diffusion process can also be described by a differential equation for a distribution function $f(x,t)$ of particles,

$$\frac{\partial f(x,t)}{\partial t} = \kappa \frac{\partial^2 f(x,t)}{\partial x^2} \ , \tag{2.45}$$

which can describe heat transport, diffusion of gases, or magnetic diffusivity on solar and stellar surfaces (i.e., manifested as meridional flows during a solar/stellar activity cycle).

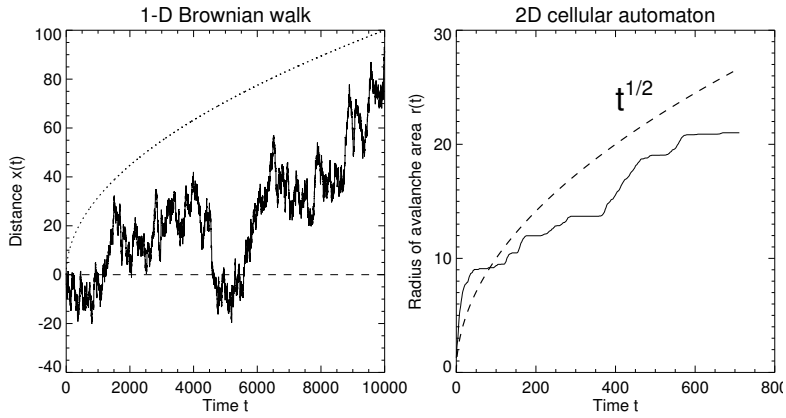


Fig. 2.11: Numerical simulation of 1D Brownian random walk (left side) and size increase of a 2D cellular automaton avalanche (right side) as a function of time t. Both processes follow the trend $x(t) \propto t^{1/2}$ expected for classical diffusion.

An example of such a diffusive random walk is simulated in Fig. 2.11 (left panel) for the 1-D variable $x(t)$. We have to be aware that a diffusive random walk $x(t)$ of a single particle can show large deviations from the expected trend $x(t) \propto t^{1/2}$, which is only an expected behavior for the statistical mean of many random walks $\langle x(t) \rangle = (1/N) \sum_i^N x_i(t)$. In Fig. 2.11 (right panel) we show also the time evolution of the mean radius $\langle r(t) \rangle$ of a simulated cellular automaton avalanche (taken from Fig. 2.8, bottom right), which shows the same trend of a time-dependence of $r(t) \propto t^{1/2}$. Apparently, the enveloping volume of unstable cells in a SOC avalanche, defined by the a mathematical re-distribution rule applied to a coarse-grained lattice with a rough surface of stable, meta-stable, and unstable nodes produces a similar time evolution as a random walk (as visualized in Fig. 2.6), leading us to the fractal-diffusive SOC model described in Section 2.2.2. This behavior of a diffusive transport process is therefore common to both the Brownian motion (or classical diffusion) and to a cellular automaton SOC process, but it operates in a SOC process only during a finite time interval, namely as long as avalanche propagation is enabled by finding unstable nearest neighbors, while a classical diffusion process goes on forever without stopping. Therefore, we cannot define an event and occurrence frequency distributions. However, we can measure the fractal dimension of a random walk spatial pattern, which has a Hausdorff

dimension of $D_2 = 2$ in 2-D Euclidean space, or the power spectrum, i.e., $P(\nu) \propto \nu^{-2}$ (Hergarten 2002), also called *Brownian noise*.

Table 2.3: Nomenclature of noise spectra.

Power spectrum	Power index	Spectrum nomenclature
$P(\nu) \propto \nu^0$	$p = 0$	white noise
$P(\nu) \propto \nu^{-1}$	$p = 1$	pink noise, flicker noise, $1/f$ noise
$P(\nu) \propto \nu^{-2}$	$p = 2$	red noise, Brown(ian) noise
$P(\nu) \propto \nu^{-3}$	$p = 3$	black noise

Generalizations that include power spectra $P(\nu) \propto \nu^{-\beta}$ with arbitrary powerlaw exponents β have been dubbed *fractional Brownian motion (fBM)* (Hergarten 2002), for instance white noise ($\beta = 0$, where subsequent steps are uncorrelated), 1/f-noise, flicker noise, or pink noise ($\beta = 1$), Brownian noise or red noise ($\beta = 2$), or black noise ($\beta = 3$, where subsequent time steps have some strong correlations, producing slowly-varying time profiles). The latter fluctuation spectrum has been found to describe the stock market well (Cheridito 2001). Examples for these different types of fractional Brownian motion are given in Fig. 2.12, while the nomenclature of noise spectra is summarized in Table 2.3.

2.3.4 Hyper-Diffusion and Lévy Flight

The cellular automaton mechanism, one prototype of SOC dynamics, involves a mathematical re-distribution rule amongst the nearest neighbor cells (Eq. 2.1). This discretized numerical re-distribution rule has been transformed in the continuum limit to an analytical function $A(x,t)$, which can be expressed as fourth-order hyper-diffusion equation (Liu et al. 2002; Charbonneau et al. 2001),

$$\frac{\partial A}{\partial t} = -\frac{\partial^2}{\partial x^2}\kappa(A_{xx}^2)\frac{\partial^2 A}{\partial x^2} \ , \tag{2.46}$$

where A represents the placeholder of the physical quantity corresponding to the symbol $z_{i,j,k}$ in Eq. 2.1, x is the S-dimensional space coordinate ($x_{i,j,k}$ for $S = 3$), with second-order centered differencing in space, forward differencing in time, and $\kappa = 1/(2S)$ the hyper-diffusion coefficient for Euclidean dimension S. However, in contrast to classical physics, the hyper-diffusion coefficient κ is subject to a threshold value in the SOC model,

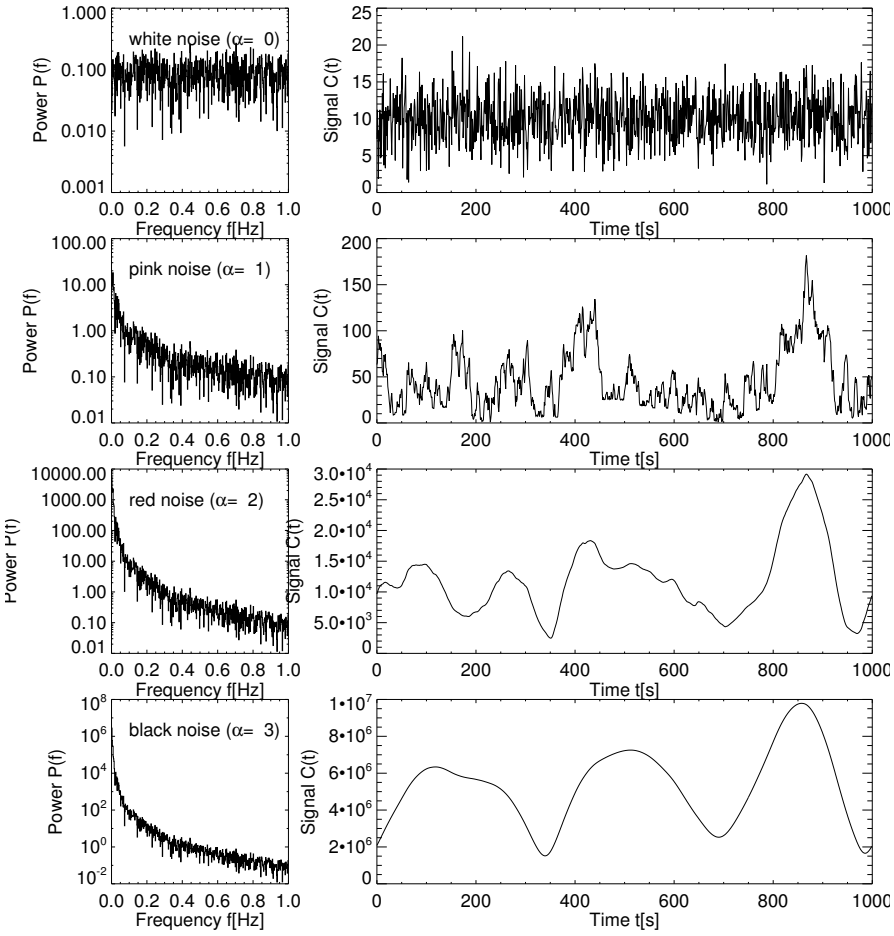


Fig. 2.12: Noise power spectra (left panels) and corresponding time series (right panels) for power spectral indices $p = 0$ (top row: white noise spectrum), $p = 1$ (second row: pink noise spectrum), $p = 2$ (third row: red noise spectrum), and $p = 3$ (bottom row: black noise spectrum). The white noise spectrum is multiplied with ν^{-p} in the other cases. The time series are reconstructed with the inverse Fourier transform (Aschwanden 2011a).

$$\kappa = \begin{cases} \kappa_a & \text{if} \quad \Delta A^2 > A^2_{\rm crit} \\ 0 & \text{otherwise} \end{cases} \quad . \tag{2.47}$$

So, on a basic level, the evolution of a cellular automaton avalanche translates into a hyper-diffusion process, as demonstrated in Liu et al. (2002). How can it be described by classical diffusion as we discussed in the foregoing Section and in the fractal-diffusive SOC model in Section 2.2.2? The major differ-

ence of the two apparently contradicting descriptions lies in the fractality: The fourth-order hyper-diffusion system (subjected to a threshold instability in the diffusion coefficient κ) is applied to a Euclidean space with dimension S, while the fractal-diffusive SOC model has a second-order (classical) diffusion coefficient in a fractal volume $V_S \propto x^{D_S}$ with a fractal dimension $D_S \approx (1+S)/2$. The two diffusion descriptions apparently produce an equivalent statistics of avalanche volumes V_S, as demonstrated by the numerical simulations of powerlaw distribution functions $N(V_S)$ by both models (Liu et al. 2002).

Other modifications of random walk or diffusion processes have been defined in terms of the probability distribution function of step sizes. Classical diffusion has a normal (Gaussian) distribution of step sizes, which was coined *Rayleigh flight* by Benoit Mandelbrot, while *Lévy flight* was used for a heavy-tailed probability distribution function (after the French mathematician Paul Pierre Lévy). Related are also the heavy-tailed Pareto probability distribution functions (Arnold 1983). Lévy flight processes include occasional large-step fluctuations on top of classical diffusion, which is found in earthquake data, financial mathematics, cryptography, signal analysis, astrophysics, biophysics, and solid state physics. A number of SOC phenomena have been analyzed in terms of Lévy flight processes, such as rice piles (Boguna and Corral 1997), random walks in fractal environments (Hopcraft et al. 1999; Isliker and Vlahos 2003), solar flare waiting time distributions (Lepreti et al. 2001), or extreme fluctuations in the solar wind (Moloney and Davidsen 2010, 2011).

2.3.5 Nonextensive Tsallis Entropy

Related to Lévy flight is also the *nonextensive Tsallis entropy*, which originates from the standard extensive *Boltzmann-Gibbs-Shannon (BGS)* statistics, where the entropy S is defined as,

$$S = -k_B \sum p_i \ln p_i \ , \tag{2.48}$$

with k_B the Boltzmann constant and p_i the probabilities associated with the microscopic configurations. The standard BGS statistics is called *extensive* when the correlations within the system are essentially local (i.e., via nearest-neighbor interactions), while it is called *nonextensive* in the case when they are non-local or have long-range coupling, similar to the local correlations in classical diffusion and non-local steps in the Lévy flights. As an example, the dynamic complexity of magnetospheric substorms and solar flares has been described with nonextensive Tsallis entropy by Balasis et al. (2011). Although nonlocal interactions are not part of the classical BTW automaton model, the extensive Tsallis entropy produces statistical distributions with

similar fractality and intermittency. The nonextensive Tsallis entropy has a variable parameter q,

$$S_q = k\frac{1}{q-1}\left(1 - \sum_{i=1}^{W} p_i^q\right) \ , \tag{2.49}$$

where q is a measure of the nonextensivity of the system. A value of $q = 1$ corresponds to the standard extensive BGS statistic, while a larger value $q > 1$ quantifies the non-extensivity or importance of long-range coupling. A value of $q = 1.84$ was found to fit data of magnetospheric substorms and solar flare soft X-rays (Balasis et al. 2011).

2.3.6 Turbulence

Turbulence is probably the most debated contender of SOC processes (e.g., Dmitruk and Gomez 1997; Dmitruk et al. 1998; Galtier and Pouquet 1998; Galtier 1999, 2001; Einaudi and Velli 1999; Buchlin et al. 2005; Boffetta et al. 1999), because of the many common observational signatures, such as the (scale-free) powerlaw distributions of spatial and temporal scales, the power spectra of time profiles, random waiting time distributions, spatial fractality, and temporal intermittency.

Let us first define turbulence. Turbulence was first defined in fluid dynamics in terms of the Navier-Stokes equation, and then by the theory of Kolmogorov (1941). A fluid is laminar at low Reynolds numbers (say at Reynolds numbers of $R \lesssim 5000$, defined by the dimensionless ratio of inertial to viscous forces, i.e., $R = vL/\nu_{\rm visc}$ with v the mean velocity of an object relative to the fluid, L the characteristic linear dimension of the fluid, and ν_{vis} the kinematic viscosity), while it becomes turbulent at high Reynolds numbers.

SOC phenomena and turbulence have been studied in astrophysical plasmas (such as in the solar corona, in the solar wind, or in stellar coronae), where the magneto-hydrodynamic (MHD) behavior of the plasma is expressed by the continuity equation, the momentum equation and induction equation,

$$\rho\frac{D\mathbf{v}}{Dt} = -\nabla p - \rho\mathbf{g} + (\mathbf{j}\times\mathbf{B}) + \nu_{\rm visc}\rho\left[\nabla^2\mathbf{v} + \frac{1}{3}\nabla(\nabla\cdot\mathbf{v})\right] \ , \tag{2.50}$$

$$\frac{\partial\mathbf{B}}{\partial t} = \nabla\times(\mathbf{v}\times\mathbf{B}) + \eta_m\nabla^2\mathbf{B} \ , \tag{2.51}$$

with ρ being the plasma density, p the pressure, $\mathbf{B}$ the magnetic field, $\mathbf{j}$ the electric current density, $\nu_{\rm visc}$ the kinematic viscosity or shear viscosity, $\eta_m = c^2/4\pi\sigma$ the magnetic diffusivity, and σ the electric conductivity. In the solar corona and in the solar wind, the magnetic Reynolds number is sufficiently high ($R_m = vL/\eta_m \approx 10^8 - 10^{12}$) to develop turbulence. Turbulence

in the coronal plasma or in the solar wind is created by initial large-scale disturbances (for instance by the convective random motion in sub-photospheric layers or by the shock-like expansion of flares and coronal mass ejections). The large-scale disturbances pump energy into the coronal magnetic field or heliospheric solar wind at large scales, which cascade in the case of turbulent flow to smaller scales, where the energy can be more efficiently dissipated by friction, which is quantified by the kinematic or (Braginskii) shear viscosity coefficient $\nu_{\rm visc}$. Ultimately, the energy of an MHD turbulent cascade is dissipated in the solar wind at the spatial scale of (gyrating) thermal protons (≈ 50 km), and at the scale of thermal electrons (≈ 0.5 km). The resulting power spectrum of the solar wind shows a power spectrum of $P(\nu) \propto \nu^{-5/3}$ at frequencies below the proton scale, $P(\nu) \propto \nu^{-7/3}$ between the proton and electron scale, and $P(\nu) \propto \nu^{-4}$ beyond the electron scale (Fig. 2.13). For literature references on MHD turbulence in the solar corona and in the solar wind see, e.g., Aschwanden (2011a, Section 10.4 therein).

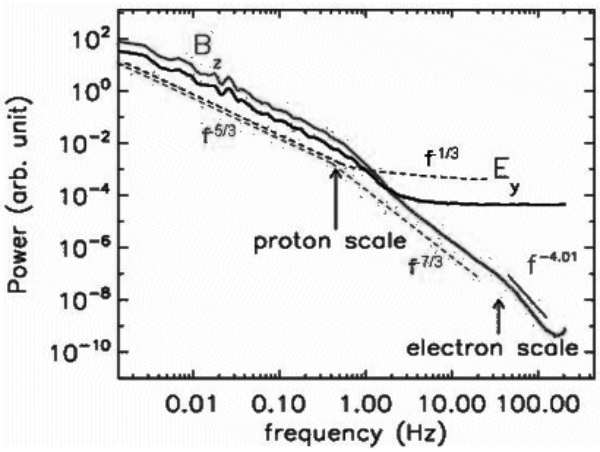


Fig. 2.13: A spectrum of the solar wind is shown, based on *CLUSTER* observations from large to small scales, with the proton and electron gyroradius scale indicated. The solar wind spectrum is interpreted in terms of a turbulent MHD cascade, with the theoretically predicted slopes of $f^{-5/3}$ and $f^{-7/3}$ from gyro-kinetic theory. The plot proves that the energy continues cascading below the proton scale down to the electron scale, where it is converted to heat (via electron Landau damping resonance) causing the steepening of the B_z spectrum to f^{-4} (Howes et al. 2008; Sahraoui et al. 2009; credit: ESA, CLUSTER).

What is the relationship between turbulence and SOC processes? The transition from laminar flow to turbulent flow at a critical Reynolds number $R_{\rm crit}$ represents a similar thresholded instability criterion as the critical value $z_{\rm crit}$ in SOC systems (although there are two different schools of definin-

ing "critical", either in terms of a global control parameter such as used in Ising models, or in terms of a local instability threshold for transport processes as used in SOC sandpiles). Above the critical threshold for turbulent flows, a turbulent avalanche can occur with subsequent cascading from vortices at large scales and little energy dissipation towards smaller scales with stronger energy dissipation. Turbulent energy dissipation probably reduces the mean velocity of the fluid particles that are faster than the background flow and laminar flow could be restored with lower particle velocities and a lower Reynolds number. To make this process self-organizing, we need also a driver mechanism that brings the fluid speed back up to the critical Reynolds number. In the solar wind, for instance, there is systematic acceleration with heliocentric distance, which could drive the system from laminar back to turbulent flows, and thus it could be self-organizing. Fluctuations of the solar wind speed could therefore be considered as SOC avalanches. This could explain the scale-free powerlaw distributions of spatial, temporal, and energy scales measured in solar wind fluctuations, its fractality and intermittency. If the critical threshold is exceeded only in localized regions, rather than in the entire system in a fully turbulent state, energy dissipation would also occur in locally unstable regions, similar to SOC avalanches, and thus the two processes may exhibit the same statistical distributions. Consequently, turbulence could qualify as a SOC process if it is driven near the laminar-turbulent critical Reynolds number in a self-organizing way. However, fully developed turbulence, where the entire system is governed by a high (supercritical) Reynolds number, would correspond to a fast-driven SOC system in a catastrophic phase with permanent avalanching, without restoring the critical state as in a slowly-driven SOC system.

2.3.7 Percolation

Percolation[1] controls a transport process that depends on the connectedness and propagation probability of nearest-neighbor elements, and has a lot in common with diffusion, fractal structures, and SOC avalanches. A classical paradigm of a percolation process is coffee percolation, which contains a solvent (water), a permeable substance (coffee grounds), and soluble constituents (aromatic chemicals). The bimodal behavior of a percolation process can be best expressed with the observation whether a liquid that is poured on top of some porous material, will be able to make its way from hole to hole and reach the bottom. The answer will very much depend on the inhomogeneity characteristics of the porous material. In the subcritical state, a percolating cluster will exponentially die out, while it will propagate all the way to the bottom in a supercritical state. The process can be described by

[1] See also Chapter 4 in this book by Alexander Milovanov on percolation.

directed percolation, where the connections between two nearest neighbors can be open and let the liquid pass with probability p, or they can be closed and the probability to pass is $(1-p)$. Completion of a pass from the top to the bottom can be expressed as the combined statistical probability along the entire pass. The critical value $p_{\rm crit}$ that decides between the two outcomes is $p_{\rm crit} = 1/2$ for a 2-D bond process. So, this process has an extreme sensitivity to the system probability p, depending whether it is below or above the critical value. In SOC systems, the outcome of an avalanche is independent of the initial conditions of the triggering disturbance. Also, the material in a percolation process has a constant value of p and does not self-adjust to a critical value, and thus it does not behave like a SOC system. However, what a percolation process has in common with a SOC system is the fractality and intermittency of propagating features. Other applications of percolation theory are in physics, material science, complex networks, epidemiology, and geography.

2.3.8 Phase Transitions

In classical thermodynamics, a phase transition describes the transformation of one phase to another state of matter (solid, liquid, gas, plasma). For instance, transitions between solid and liquid states are called "freezing" and "melting", between liquid and gas are "vaporization" and "condensation", between gas and plasma are "ionization" and "deionization". Phase transitions from one state to another can be induced by changing the temperature and/or pressure.

Since there are critical values of temperature and pressure that demarcate the transitions, such as the freezing temperature at 0^0 C or the boiling point of 100^0 C, we may compare these critical values with the critical threshold $z_{\rm crit}$ in SOC systems. Could this critical value be restored in a self-organizing way? To some extent, the daily weather changes can be self-organizing. Let us assume some intermediate latitude on our planet where the temperature is around the freezing point at some seasonal period, say Colorado or Switzerland around November. The temperature may drop below the freezing point during night, triggering snowfall, and may raise slightly above freezing point during sunny days, triggering melting of snow. In this case, the day-night cycle or the Earth's rotation causes the temperature to oscillate around the freezing point, and therefore behaves like a self-organizing system. Additional variation of cloud cover may introduce fluctuations on even shorter time scales. However, this bimodal behavior of the temperature is not a robust operation mode of self-organized criticality, because seasonal changes will bring the average daily temperature systematically out of the critical range around the freezing point. So, such temporary fluctuations around a critical point seem to be consistent with a SOC system only on a temporary basis or in inter-

mittent (seasonal) time intervals. A similar concept of sweeping through a critical point as an alternative to SOC to get powerlaws without parameter tuning was also proposed by Sornette (1994).

In fact, a number of metereological phenomena have been found to exhibit powerlaw distributions and where associated with SOC behavior, such as Nile river fluctuations (Hurst 1951), rainfall (Andrade et al. 1998), cloud formation (Nagel and Raschke 1992), climate fluctuations (Grieger 1992), aerosols in the atmosphere (Kopnin et al. 2004), or forest fires (Kasischke and French 1995, Malamud et al. 1998). Phase transitions may be involved in some of these processes (rainfall, cloud formation, forest fires).

Phase transitions are also involved in some laboratory experiments in material or solid-state physics, quantum mechanics, and plasma physics, that have been associated with SOC behavior, if we extend the term phase transition also to morphological changes between highly-ordered and/or chaotic-structured patterns. Such phase transitions with SOC behavior have been found in transitions between the ferromagnetic and paramagnetic phases, in specimens of ferromagnetic materials settling into one of a large number of metastable states that are not necessarily the energetically lowest ones (Che and Suhl 1990), in avalanche-like topological rearrangements of cellular domain patterns in magnetic garnet films (Babcock and Westervelt 1990), in the noise in the magnetic output of a ferromagnet when the magnetizing force applied to it is charged (Cote and Meisel 1991), called the Barkhausen effect, in superconducting vortex avalanches in the Bean state (Field et al. 1995), in tokamak plasma confinement near marginal stability driven by turbulence with a subcritical resistive pressure gradient (Carreras et al. 1996), or in the electrostatic floating potential fluctuations of a DC glow charge plasma (Nurujjaman and Sekar-Iyenbgar 2007).

In summary, many SOC phenomena are observed during phase transitions, producing avalanche-like rearrangements of geometric patterns, but a phase-transition is not necessarily self-organizing in the sense that the critical control parameter stays automatically tuned to the critical point. A sandpile maintains its critical slope automatically in a slowly-driven operation mode. If a thermostat would be available, temperature-controlled phase transitions could be in a self-organized operation mode, which indeed can be arranged in many situations. So we conclude that phase transition with some kind of "thermostat" tuned to the critical point produce a similar behavior as SOC systems, but this conclusion can be disputed if the term "critical" is used in a more restricive sense as originally defined in the context of Ising models (Pruessner and Peters 2006, 2008; Alava et al. 2008).

2.3.9 Network Systems

The prototype of a SOC cellular automaton model is a regular lattice grid, where re-distributions occur with nearest neighbors cells. In contrast, networks are irregular nets of nodes that are inter-connected in manifold patterns, containing not necessarily only nearest-neighbor connections, but also arbitrary non-local, long-range connections, (for an introduction into networks see, e.g., Newman 2010). Popular examples of networks are subway maps, city maps, road maps, airplane route maps, electric grid maps, social networks, company organizational charts, financial networks, etc, all being created by human beings in some way or another. One big difference to SOC systems is therefore immediately clear, the addition of long-range connections. Furthermore, while the statistical probabilities for nearest-neighbor interactions are identical for each lattice point in a SOC cellular automaton model, networks can have extremely different connection probabilities at each node point. Nodes that have most of the connections are also called hubs. For instance, the American airline *United* has 378 destinations, but serves most of them from 10 hubs (Denver, Washington DC, San Francisco, Los Angeles, Chicago, Houston, Cleveland, Newark, Narita, Guam) (Fig. 2.14). Since the connection probability varies greatly from node to node, we do expect different clustering patterns in networks and in SOC lattice systems.

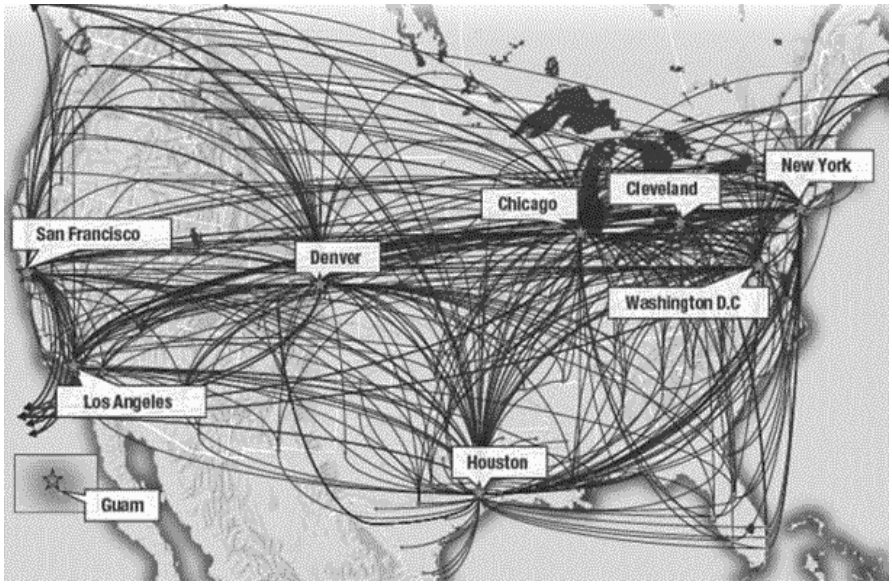


Fig. 2.14: Network of American airline *UNITED* with 8 US-based hubs and connections to 378 destinations in 59 countries (Credit: United).

Nevertheless, SOC behavior has been studied in network systems like city growth and urban growth (Zipf 1949; Zanette 2007), cotton prizes (Mandelbrot 1963), financial stock market (Scheinkman and Woodford 1994; Sornette et al. 1996; Feigenbaum 2003; Bartolozzi et al. 2005), traffic jams (Bak 1996; Nagel and Paczuski 1995), war casualties (Richardson 1941, Levy 1983), social networks (Newman et al. 2002), internet traffic (Willinger et al. 2002), and language (Zipf 1949). SOC phenomena with powerlaw-like frequency distributions have been observed for phenomena that occur in these networks, measured by some quantity that expresses fluctuations above some threshold or noise level. Spatial patterns of network phenomena have been found to be fractal, and time profiles of these fluctuations were found to be intermittent, all hallmarks of SOC systems. So, are those network phenomena self-organizing? What is the critical threshold and what drives the system back to the critical threshold? The driving force of cotton prices, lottery wins, and stockmarket fluctuations is simply the human desire to gain money and to make profit. The driving mechanism for traffic jams is certainly the daily need for commuting and transportation in a finite road network. The unpredictable number of casualties in crimes and wars is a bit more controversial, but nobody would disagree that the human desire for power and control plays a role. The driving mechanism for social networks is most likely linked to curiosity and the desire for information. Since all these human traits are quite genuine and persistent, a self-organizing behavior is warranted, where the system is continuously driven towards a critical state, with occasional larger fluctuations (Wall Street crash, Second World War, or other "social catastrophes"), on top of the daily fluctuations on smaller scales. In summary, most network-related phenomena can exhibit SOC behavior, apparently with similar powerlaw-like statistics in grids with non-uniform connectivities and non-local long-range connections (in network systems) as in regular lattice grids with nearest-neighbor interactions (in classical SOC systems).

2.3.10 Chaotic Systems

Chaotic systems are nonlinear dissipative dynamical processes that are in principle deterministic, but exhibit "chaotic" behavior in the sense of irregular and fractal geometry, and intermittent time evolution. Some chaotic systems can be described as simple as with two (such as the Lotka-Volterra equation) or three coupled differential equations (e.g., the Lorenz equations). Classical examples include forced pendulum, fluids near the onset of turbulence, lasers, nonlinear optical devices, Josephson junctions, chemical reactions, the three-body system in celestial mechanics, ecological population dynamics, or the heartbeat of biological organisms (for an overview see, e.g., Schuster 1988, and references therein). If two dynamical variables (say $x(t)$ and $y(t)$) are plotted in phase space (y versus x), chaotic systems often ex-

hibit a cyclic behavior, with a limit cycle that is also called *strange attractor*. If a dynamical system is driven from small disturbances around an equilibrium solution, where it exhibits a quasi-linear behavior, bifurcations of system variables $x(t)$ can occur at critical values with transitions to chaotic and intermittent behavior (also called route to chaos, e.g., Shraiman et al. 1981). Such bifurcations ressemble the phase transitions we discussed earlier (Section 2.3.8), which we generalized to transitions from highly-ordered to chaotic-structured patterns. The complexity of nonlinear chaotic systems is often characterized with the *dimension of strange attractors* (e.g., Grassberger 1985; Guckenheimer 1985), which approximately corresponds to the number of coupled differential equations that is needed to describe the system dynamics.

If we compare the dynamics of chaotic systems with SOC processes, a major difference is the deterministic evolution of chaotic systems, while SOC avalanches are triggered statistically independently of each other, even when they occur subsequently within a small (waiting) time interval (but see Welinder et al. 2007 for a counter example). A chaotic system evolves with a system-wide dynamical behavior that is influenced by the entire system, as the coupled differential equations express it, which includes nearest neighbors as well as non-local, long-range coupling, while an avalanche in a SOC system evolves as a result of nearest-neighbor interactions. Do chaotic systems have critical thresholds that are self-organizing? Chaotic systems do have critical points, where onset of chaos starts, like at pitchfork bifurcation point that leads to frequency doubling, but there is usually not a driver that keeps the system at this critical point. For instance, deterministic chaos sets in for fluids at the transition from the laminar to the turbulent regime, but generally there exists no driving mechanism that keeps a chaotic system at the critical Reynolds number, so that the system is self-organizing (Section 2.3.6). However, in other respects, chaotic systems exhibit similar complexity as SOC systems, regarding fractality and intermittency, and even powerlaw distributions may result in the statistics of chaotic fluctuations. For weakly nonlinear dissipative systems near the limit cycle, however, small fluctuations around a specific spatial scale or temporal scale may produce Gaussian-like distributions, e.g., time scales may be centered around the inverse frequency of the limit cycle.

Let us mention some examples of chaotic behavior in astrophysics. Time series analysis of the X-ray variability of the neutron star Her X-1 revealed a low-dimensional attractor (Voges et al. 1987; Cannizzo et al. 1990), as well as for the Vela pulsar (Harding et al. 1990). Transient chaos was detected for the low-mass X-ray binary star Scorpius X-1 (Scargle et al. 1993; Young and Scargle 1996), as well as for the R scuti star (Buchler et al. 1996). The N-body system in celestial mechanics can lead to chaotic behavior, such as for the Saturnian moon Hyperion (Boyd et al. 1994). Chaotic behavior with low-dimensional strange attractors and transitions to period doubling was also found in solar radio bursts with quasi-periodic time series (Kurths and

	Powerlaw of spatial scales N(L)	Powerlaw of time scales N(T)	Powerlaw of total energy N(E)	Powerlaw of energy dissipation rate N(F),N(P)	Fractal geometry	Intermittency in time evolution	Statistically independent trigger of events	Critical threshold restoration	Nearest-neighbor interactions	Non-local (long-range) coupling
Self-Organized Criticality (SOC)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Exponential-Growth model (EG-SOC)	Y	N	Y	Y	N	N	Y	Y	Y	-
Fractal-Diffusive model (FD-SOC)	Y	Y	Y	Y	Y	Y	Y	Y	Y	-
Forced Criticality model (FSOC)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Self-Organization (SO)	N	-	-	-	N	N	N	N	Y	Y
Brownian Motion, classical diffusion	Y	-	-	-	Y	N	N	N	Y	Y
Hyper-Diffusion (thresholded)	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Levy Flight (thresholded)	Y	?	?	?	Y	Y	Y	N	Y	Y
Nonextensive Tsallis entropy	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Turbulence (laminar/turbulent)	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Turbulence (fully turbulent)	Y	?	?	?	Y	Y	N	N	Y	Y
Percolation	Y	?	?	?	Y	Y	N	N	Y	N
Phase Transitions	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Phase Transitions (with “thermostats’)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Network systems	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Chaotic Systems	Y	Y	Y	Y	Y	Y	N	N	Y	Y

Fig. 2.15: Matrix of SOC or SOC-like processes (rows) and observational properties (columns), with an evaluation whether a specific process can exhibit a particular observational characteristics (Y=Yes, N=No, and the symbol − is filled in for observational properties that do not apply). See Chapter 2.3 for details.

Herzel 1986, Kurths and Karlicky 1989). Chaotic dynamics was found in the solar wind (Polygiannakis and Moussas 1994), in hydrodynamic convection simulations of the solar dynamo (Kurths and Brandenburg 1991), and in solar cycle observations (Kremliovsky 1994; Charbonneau 2001; Spiegel 2009). Again, all these examples reflect a system-wide nonlinear behavior, where subsequent fluctuations are highly correlated in space and time, unlike the

statistically independent trigger of SOC avalanches. However, chaotic systems may exhibit similar statistics of fractal spatial and intermittent time scales at critical points and transitions to chaotic dynamics, without having an intrinsic mechanism that keeps the chaotic system near this critical point in a self-organizing way.

2.3.11 Synopsis

At the end of this chapter we summarize the characteristics of SOC, SOC-related, and non-SOC processes in a matrix as shown in Fig. 2.15. The SOC or SOC-like processes are listed in the rows of Fig. 2.15, while the observational characteristics are listed in the columns. The matrix tabulates whether a specific SOC-like process can exhibit the main properties of SOC processes, such as the powerlaw distributions of various parameters, the fractal geometry, the temporal intermittency, the statistical independence of events, the restoration of a critical threshold, nearest-neighbor interactions, and non-local or long-range coupling. The properties listed in Fig. 2.15 reflect general trends rather than strictly-valid matches. Many processes can exhibit powerlaw distributions of parameters, but there exist always exceptions or deviations from strict powerlaw distributions (such as due to finite-size effects, sensitivity-limited sampling, exponential fall-off in cumulative distributions, nonstationary drivers, inhomogeneous media, etc.) The probably most fundamental characteristics of SOC processes is a suitable mechanism that restores the critical threshold for a instability, which often is not automatically operating (such as in turbulence, phase transitions, or chaotic systems), but can be artificially or naturally added to a process (such as a "thermostat" for phase transitions). A non-discriminant criterion between SOC and non-SOC processes is the presence of short-range and long-range coupling, which appears to exist for classical SOC cellular automaton mechanisms as well as for other processes (e.g., turbulence, phase transitions, network systems, and chaotic systems). Finally, fractality, intermittency, and powerlaw behavior is present in most of the processes also, so it is not a good discrimination criterion between SOC and non-SOC processes, although powerlaws have always been considered as the hallmark of SOC processes.

Acknowledgements: This work has benefitted from fruitful discussions with Henrik Jensen, Nicholas Watkins, Jürgen Kurths, Lawrence Schulman, Gabriela Monacu, Ewan Cameron, Jonathan Katz, comments from two anonymous referees, and from the *International Space Science Institute (ISSI)* at Bern Switzerland, which hosted and supported a workshop on *Self-Organized Criticality and Turbulence* during October 15-19, 2012.

2.4 References

Alava, M.J., Laurson, L, Vespignani, A., and Zapperi, S. 2008, *Comment on "Self-organized criticality and absorbing states: Lessons from the Ising model*, Phys. Rev. E. **77**(4), 048101.

Andrade, R.F.S., Schellnhuber, H.J., and Claussen, M. 1998, *Analysis of rainfall records: possible relation to self-organized criticality*, Physica A **254**(3/4), 557-568.

Angelopoulos, V., Coroniti, F.V., Kennel, C.F., Kivelson, M.G., Walker, R.J., Russell, C.T., McPherron, R.L., Sanchez, E., Meng, C.I., Baumjohann, W., Reeves, G.D., Belian, R.D., Sato, N., Friis-Christensen, E., Sutcliffe, P.R., Yumoto,K., Harris,T. 1996, *Multipoint analysis of a bursty bulk flow event on April 11, 1985*, J. Geophys. Res. **101**(A3), 4967-4990.

Angelopoulos, V., Mukai, T., and Kokubun, S. 1999, *Evidence for intermittency in Earth's plasma sheet and implications for self-organized criticality*, Phys. Plasmas **6**(11), 4161-4168.

Antipov, S.V., Nezlin, M.V., Snezhkin, E.N., and Trubnikov, A.S. 1985, *The Rossby autosoliton and a laboratory model of the Jupiter Red SPOT*, Zhurnal Eksperimental noi i Teoreticheskoi Fiziki, **89**, 1905.

Arnold, B.C. 1983, *Pareto Distributions*, International Co-operative Publishing House. ISBN 0-89974-012-X.

Aschwanden, M.J., Dennis, B.R., and Benz, A.O. 1998, *Logistic avalanche processes, elementary time structures, and frequency distributions of flares*, Astrophys. J. **497**, 972-993.

Aschwanden, M.J. 2004, *Physics of the Solar Corona. An Introduction*, ISBN 3-540-22321-5, PRAXIS, Chichester, UK, and Springer, Berlin, 842p.

Aschwanden, M.J. 2011a, *Self-Organized Criticality in Astrophysics. The Statistics of Nonlinear Processes in the Universe*, ISBN 978-3-642-15000-5, Springer-Praxis: New York, 416p.

Aschwanden, M.J. 2011b, *The state of self-organized criticality of the Sun during the last 3 solar cycles. I. Observations*, Solar Phys. **274**, 99-117.

Aschwanden, M.J. 2012, *A statistical fractal-diffusive avalanche model of a slowly-driven self-organized criticality system*, Astron.Astrophys. **539**, A2 (15 p).

Aschwanden, M.J. and Freeland, S.L. 2012, *Automated solar flare statistics in soft X-rays over 37 years of GOES observations - The Invariance of self-organized criticality during three solar cycles*, Astrophys. J. **754**, 112.

Babcock, K.L. and Westervelt, R.M. 1990, *Avalanches and self-organization in cellular magnetic-domain patterns*, Phys. Rev. Lett. **64**(18), 2168-2171.

Bagnold, R. (1941) 2005, *The Physics of Blown Sand and Desert Dunes*, Dover Publications. ISBN 978-0-486-43931-0.

Bak, P., Tang, C., and Wiesenfeld, K. 1987, *Self-organized criticality - An explanation of 1/f noise*, Phys. Rev. Lett. **59**(27), 381-384.

Bak, P., Tang, C., and Wiesenfeld, K. 1988, *Self-organized criticality*, Phys. Rev. A **38**(1), 364-374.

Bak, P. and Chen, K. 1989, *The physics of fractals*, Physica D **38**, 5-12.

Bak, P., Chen, K. and Creutz, M. 1989, *Self-organized criticality in the "Game of Life"*, Nature **342**, 780-781.

Bak, P. and Sneppen, K. 1993, *Punctuated equilibrium and criticality in a simple model of evolution*, Phys. Rev. Lett. **71**(24), 4083-4086.

Bak, P. 1996, *How nature works*, Copernicus, Springer-Verlag, New York.

Balasis, G., Daglis, I.A., Anastasiadis, A., Papadimitriou, C., Mandea, M., and Eftaxias,K. 2011, *Universality in solar flare, magnetic storm and earthquake dynamics using Tsallis statistical mechanics*, Physica A **390**, 341-346.

Bartolozzi, M., Leinweber, D.B., and Thomas, A.W. 2005, *Self-organized criticality and stock market dynamics: an empirical study*, Physica A **350**, 451-465.

Beltrami,E. 1987, *Mathematics for Dynamic Modeling*, New York: Academic Press Inc.

Boffetta, G., Carbone, V., Giuliani, P., Veltri, P., and Vulpiani, A. 1999, *Power laws in solar flares: self-organized criticality or turbulence*, Phys. Rev. Lett. **83**(2), 4662-4665.

Boguna, M. and Corral, A. 1997, *Long-tailed trapping times and Lévy flights in a self-organized critical granular system*, Phys.Rev.Lett. **78**(26), 4950-4953.

Boyd, P.T., Mindlin, G.B., Gilmore, R., and Solari, H.G. 1994, *Topological analysis of chaotic orbits: revisiting Hyperion*, Astrophys. J. **431**, 425-431.

Buchler, J.R., Kollath, Z., Serre, T., and Mattei, J. 1996, *Nonlinear analysis of the light curve of the variable star R Scuti*, Astrophys. J. **462**, 489-501.

Buchlin, E., Galtier, S., and Velli, M. 2005, *Influence of the definition of dissipative events on their statistics*, Astron. Astrophys. **436**, 355-362.

Burridge R. and Knopoff L. 1967, *Model and theoretical seismicity*, Seis. Soc. Am. Bull. **57**, 341-347.

Camazine,S., Deneubourg,J.L., Frank,N.R., Sneyd,J., Theraulaz,G., and Bonabeau,E. 2001, *Self-Organization in Biological Systems*, Princeton University Press.

Cannizzo, J.K., Goodings, D.A., and Mattei, J.A. 1990, *A search for chaotic behavior in the light curves of three long-term variables*, Astrophys. J. **357**, 235-242.

Carreras, B.A., Newman, D., Lynch, V.E., and Diamond, P.H., 1996, *A model realization of self-organized criticality for plasma confinement*, Phys. Plasmas **3**(8), 2903-2911.

Chang, T.S. 1992, *Low-dimensional behavior and symmetry breaking of stochastic systems near criticality - Can these effects be observed in space and in the laboratory*, IEEE Trans. Plasma Sci. **20**(6), 691-694.

Chang, T.S. 1998a, *Sporadic, Localized reconnections and mnultiscale intermittent turbulence in the magnetotail*, in *Geospace Mass and Energy Flow* (eds. Horwitz, J.L., Gallagher, D.L., and Peterson, W.K.), AGU Geophysical Monograph **104**, p.193.

Chang, T.S. 1998b, *Multiscale intermittent turbulence in the magnetotail*, in *Proc. 4th Intern. Conf. on Substorms*, (eds. Kamide, Y. et al.), Kluwer Academic Publishers, Dordrecht, and Terra Scientific Company, Tokyo, p.431.

Chang, T.S. 1999a, *Self-organized criticality, multi-fractal spectra, and intermittent merging of coherent structures in the magnetotail*, Astrophys. Space Sci. **264**, 303-316.

Chang, T.S. 1999b, *Self-organized criticality, multi-fractal spectra, sporadic localized reconnections and intermittent turbulence in the magnetotail*, Phys. Plasmas **6**(11), 4137-4145.

Charbonneau, P., McIntosh, S.W., Liu, H.L., and Bogdan, T.J. 2001, *Avalanche models for solar flares*, Solar Phys. **203**, 321-353.

Che, X. and Suhl, H., 1990, *Magnetic domain pattern as self-organizing critical systems*, Phys. Rev. Lett. **64**(14), 1670-1673.

Cheridito, P. 2001, *Regularizing fradtional Brownian motion with a view towards stock price modelling*, PhD dissertation, ETH Zürich.

Consolini, G. 1997, *Sandpile cellular automata and magnetospheric dynamics*, in Proc, Cosmic Physics in the year 2000, (eds. S.Aiello, N.Iucci, G.Sironi, A.Treves, and U.Villante), SIF: Bologna, Italy, Vol. **58**, 123-126.

Consolini, G. and Lui, A.T.Y. 1999, *Sign-singularity analysis of current disruption*, Geophys. Res. Lett. **26**(12), 1673-1676.

Cote, P.J. and Meisel, L.V. 1991, *Self-organized criticality and the Barkhausen effect*, Phys. Rev. Lett. **67**, 1334-1337.

Crosby, N.B., Meredith, N.P., Coates, A.J., and Iles, R.H.A. 2005, *Modelling the outer radiation belt as a complex system in a self-organised critical state*, Nonlinear Processes in Geophysics **12**, 993-1001.

Crosby, N.B. 2011, *Frequency distributions: From the Sun to the Earth*, Nonlinear Processes in Geophysics **18**(6), 791-805.

Dennis, B.R. and Zarro, D.M. 1993, *The Neupert effect: what can it tell us about the impulsive and gradual phases of solar flares*, Solar Phys. **146**, 177-190.

Dennis, B.R., Veronig, A., Schwartz, R.A., Sui, L., Tolbert, A.K., Zarro,D.M. and the RHESSI Team 2003, *The Neupert effect and new RHESSI measures of the total energy in electrons accelerated in solar flares*, Adv.Space Res. **32**, 2459.

Dmitruk, P. and Gomez, D.O. 1997, *Turbulent coronal heating and the distribution of nanoflares*, Astrophys. J. **484**, L83-L85.

Dmitruk, P., Gomez, D.O., and DeLuca, E.E. 1998, *Magnetohydrodynamic turbulence of coronal active regions and the distribution of nanoflares*, Astrophys. J. **505**, 974-983.

Einaudi, G. and Velli M. 1999, *The distribution of flares, statistics of magnetohydrodynamic turbulence and coronal heating*, Phys. Plasmas **6**(11), 4146-4153.

Feigenbaum, J.A. 2003, *Financial physics*, Rep. Prog. Phys. **66**, 1611-1649.

Fermi, E. 1949, *On the origin of the cosmic radiation*, Phys. Rev. Lett. **75**, 1169-1174.

Field, S., Witt, J., Nori, F., and Ling, X. 1995, *Superconducting vortex avalanches*, Phys. Rev. Lett. **74**, 1206-1209.

Galtier, S., and Pouquet, A. 1998, *Solar flare statistics with a one-dimensional MHD model*, Solar Phys. **179**, 141-165.

Galtier, S., 1999, *A one-dimensional magnetohydrodynamic model of solar flares: Emergence of a population of weak events, and a possible road toward nanoflares*, Astrophys. J. **521**, 483-489.

Galtier, S., 2001, *Statistical study of short quiescent times between solar flares in a 1-D MHD model*, Solar Phys. **201**, 133-136.

Getling, A.V. 1998, *Rayleigh-Bénard Convection: Structures and Dynamics*, World Scientific. ISBN 978-981-02-2657-2.

Grassberger, P. 1985, *Information aspects of strange attractors*, in *Chaos in Astrophysics* (eds. Buchler, J.R. et al.), Reidel Publishing Company: Dordrecht, p.193-222.

Grieger, B. 1992, *Quaternary climatic fluctuations as a consequence of self-organized criticality*, Physica A **191**, 51-56.

Guckenheimer J. 1985, *Clues to strange attractors*, in *Chaos in Astrophysics* (eds. Buchler, J.R. et al.), Reidel Publishing Company: Dordrecht, p.185-191.

Harding, A.K., Shinbrot, T., and Cordes, J.M. 1990, *A chaotic attractor in timing noise from the Vela pulsar?*, Astrophys. J. **353**, 588-596.

Hergarten, S. and Neugebauer, H.J. 1998, *Self-organized criticality in a sandslide model*, Geophys. Res. Lett. **25**(6), 801-804.

Hergarten, S. 2002, *Self-organized criticality in Earth systems*, Springer: New York, 272p.

Hopcraft, K.I., Jakeman, E., and Tanner, R.M.J. 1999, *Levy random walks with fluctuating step number and multiscale behavior*, Phys.Rev E **60**(5), 5327-5343.

Hoshino, M., Nishida, A., Yamamoto, T., and Kokubun, S. 1994, *Turbulent magnetic field in the distant magnetotail: bottom-up process of plasmoid formation?*, Geophys. Res. Lett. **21**(25), 2935-2938.

Howes, G.G., Dorland, W., Cowley, S.C., Hammett, G.W., Quataert, E., Schekochihin, A.A., and Tatsuno, T. 2008, *Kinetic simulations of magnetized turbulence in astrophysical plasmas*, Phys. Rev. Lett. **100**(6), 065004.

Huberman, B.A. and Adamic, L. 1999, *Growth dynamics of the World-Wide Web*, Nature **401**, 131.

Hurst, H.E. 1951, *Long-term storage capacity of reservoirs*, Trans. Am. Soc. Civil Eng. **116**, 770-799.

Isliker, H., Anastasiadis, A., Vassiliadis, D., and Vlahos, L. 1998, *Solar flare cellular automata interpreted as discretized MHD equations*, Astron. Astrophys. **335**, 1085-1092.

Isliker, H. and Vlahos, L. 2003, *Random walk through fractal environments*, Phys. Rev. E, **67**(2), id. 026413.

Jackson,E.A. 1989, *Perspectives of nonlinear dynamics*, Cambridge: Cambridge University Press

Jensen, H.J. 1998, *Self-Organized Criticality. Emergent complex behavior in physical and biological systems*, Cambridge University Press, Cambridge UK, 153 p.

Kasischke, E.S.and French N.H.F. 1995, Remote Sens. Environ. **51**, 263-275.

Katz, J.I. 1986, A model of propagating brittle failure in heterogeneous media, JGR **91**(B10), 10,412-10,420.

Kolmogorov, A. 1941, *The local structure of turbulence in incompressible viscous fluid for very large Reynold's numbers*, Doklady Akademiia Nauk SSSR **30**, 301-305.

Kopnin, S.I., Kosarev, I.N., Popel, S.I., and hyu, M.Y. 2004, *Localized structures of nanosize charged dust in Earth's middle atmosphere*, Planet. Space Sci. **52**(13), 1187-1194.

Kozelov, B.V., Uritsky, V.M., and Klimas, A.J. 2004, *Power law probability distributions of multiscale auroral dynamics from ground-based TV observations*, Geophys. Res. Lett. **31**(20), CiteID L20804.

Kremliovsky, M.N. 1994, *Can we understand time scales of solar activity?*, Solar Phys. **151**, 351-370.

Kurths, J. and Herzel, H. 1986, *Can a solar pulsation event be characterized by a low-dimensional chaotic attractor*, Solar Phys. **107**, 39-45.

Kurths, J. and Karlicky, M. 1989, *The route to chaos during a pulsation event*, Solar Phys. **119**, 399-411.

Kurths, J. and Brandenburg, A. 1991, *Lyapunov exponents for hydrodynamic convection*, Phys. Rev. A. **44**(6), 3427-3429.

Lepreti, F., Carbone, V., and Veltri, P. 2001, *Solar flare waiting time distribution: varying-rate Poisson or Levy function?*, Astrophys. J. **555**, L133-L136.

Levy, J.S. 1983, *War in the Modern Great Power System 1495-1975*, (Lexington, KY: University of Kentucky Press), p. 215.

Liu, H., Charbonneau, P., Pouquet, A., Bogdan, T., and McIntosh, S.W. 2002, *Continuum analysis of an avalanche model for solar flares*, Phys. Rev. E **66**, 056111.

Lu, E.T. and Hamilton, R.J. 1991, *Avalanches and the distribution of solar flares*, Astrophys. J. **380**, L89-L92.

Lui, A.T.Y., Chapman, S.C., Liou,K., Newell, P.T., Meng, C.I., Brittnacher, M., and Parks, G.K. 2000, *Is the dynamic magnetosphere an avalanching system?*, Geophys. Res. Lett. **27**(7), 911-914.

Malamud, B.D., Morein, G., and Turcotte D.L. 1998, *Forest fires: An example of self-organized critical behavior*, Science **281**, 1840-1842.

Mandelbrot, B.B. 1963, *The variation of certain speculative prices*, Journal of Business of the University of Chicago **36**(4), 394-419.

May, R.M. 1974, *Model Ecosystems*, Princeton: Princeton.

Mineshige, S., Ouchi,N.B., and Nishimori, H. 1994a, *On the generation of 1/f fluctuations in X-rays from black-hole objects*, Publ. Astron. Soc. Japan **46**, 97-105.

Mineshige, S., Takeuchi, M., and Nishimori, H. 1994b, *Is a black hole accretion disk in a self-organized critical state ?*, Astrophys. J. **435**, L125-L128.

Moloney, N.R. and Davidsen, J. 2010, *Extreme value statistics in the solar wind: An application to correlated Levy processes*, JGR (Space Physics), **115**, 10114.

Moloney, N.R. and Davidsen, J. 2011, *Extreme bursts in the solar wind*, GRL **381**, 14111.

Morales, L. and Charbonneau, P. 2008, *Self-organized critical model of energy release in an idealized coronal loop*, Astrophys. J. **682**, 654-666.

Nagel, K. and Raschke, E. 1992, *Self-organized criticality in cloud formation?*, Physica A **182**(4), 519-531.

Nagel, K. and Paczuski, M. 1995, *Emergent traffic jams*, Phys. Rev. E. **51**(4), 2909-2918.

Newman, M.E.J., Watts, D.J., and Strogatz, S.H. 2002, *Random graph models of social networks*, in *Self-organized complexity in the physical, biological, and social sciences*, Arthur M. Sackler Colloquia, (eds. Turcotte, D., Rundle, J., and Frauenfelder, H.), The National Academy of Sciences: Washington DC, p.2566-2572.

Newman, M.E.J. 2010, *Networks: An Introduction*, Oxford University Press.

Nurujjaman, Md. and Sekar-Iyenbgar, A.N. 2007, *Realization of SOC behavior in a DC glow discharge plasma*, Phys. Lett. A **360**, 717-721.

Pavlidou, V., Kuijpers, J., Vlahos, L., and Isliker, H. 2001, *A cellular automaton model for the magnetic activity in accretion disks*, Astron. Astrophys. **372**, 326-337.

Polygiannakis, J.M. and Moussas, X. 1994, *On experimental evidence of chaotic dynamics over short time scales in solar wind and cometary data using nonlinear prediction techniques*, Solar Phys. **151**, 341-350.

Pruessner, G. and Peters, O. 2006, *Self-organized Criticality and Absorbing States: Lessons from the Ising Model*, Phys. Rev. E **73**(2), 025106(R)-1–4.

Pruessner, G. and Peters, O. 2008, *Reply to Comment on Self-organized criticality and absorbing states: Lessons from the Ising model*, Phys. Rev. E. **77**, 048102.

Pruessner, G. 2012, *Self-organised criticality. Theory, models and characterisation*, ISBN 978-0-521-85335-4, Cambridge University Press, Cambridge.

Reed, W.J. and Hughes, B.D. 2002, *From gene families and genera to incomes and internet file sizes: Why power laws are so common in nature*, Phys. Rev. Lett. E **66**, 067103.

Richardson, L.F. 1941, *Frequency occurrence of wars and other fatal quarrels*, Nature **148**, 598.

Rosner, R., and Vaiana, G.S. 1978, *Cosmic flare transients: constraints upon models for energy storage and release derived from the event frequency distribution*, Astrophys. J. **222**, 1104-1108.

Rovinsky, A.B. and Menzinger, M. 1993, *Self-organization induces by the differential flow of activator and inhibitor*, Phys. Rev. Lett. **70**, 778-781.

Sahraoui, F., Goldstein, M.L., Robert, P., Khotyzintsev, Y.V. 2009, *Evidence of a cascade and dissipation of solar-wind turbulence at the electron gyroscale*, Phys. Rev. Lett. **102**, 231102:1-4.

Scargle, J. 1993, *Wavelet methods in astronomical time series analysis*, Internat. Conf. on *Applications of time series analysis in astronomy and meteorology*, (ed. O. Lessi), Padova, Italy.

Scheinkman, J.A. and Woodford, M. 1994, *Self-organized criticality and economic fluctuations*, Am. Econ. Rev. **84**, 417-421.

Schuster, H.G. 1988, *Deterministic Chaos: An Introduction*, Weinheim (Germany): VCH Verlagsgesellschaft, 270 p.

Shraiman, B., Wayne, C.E., and Martin, P.C. 1981, *Scaling Theory for Noisy Period-Doubling Transition to Chaos*, Phys. Rev. Lett. **46**, 935.

Sornette, D. 1994, *Sweeping of an instability: an alternative to self-organized criticality to get powerlaws without parameter tuning*, J. Phys. I (France), **4**, 209-211.

Sornette, D. 2004, *Critical phenomena in natural sciences: chaos, fractals, self-organization and disorder: concepts and tools*, Springer, Heidelberg, 528 p.

Sornette, D., Johansen, A., and Bouchard, J.P. 1996, *Stock market crashes, precursors and replicas*, J. Phys. I (France) **6**, 167-175.

Spiegel, E.A. 2009, *Chaos and Intermittency in the Solar Cycle*, Space Sci. Rev. **144**, 25-51.

Takalo, J. 1993, *Correlation dimension of AE data*, Ph. Lic. Thesis, *Laboratory report 3*, Dept. Physics, University of Jyväskylä.

Takalo, J., Timonem, J., Klimas, A., Valdivia, J., and Vassiliadis, D. 1999a, *Nonlinear energy dissipation in a cellular automaton magnetotail field model* Geophys. Res. Lett. **26**(13), 1813-1816.

Takalo, J., Timonem, J., Klimas, A., Valdivia, J., and Vassiliadis, D. 1999b, *A coupled-map model for the magnetotail current sheet*, Geophys. Res. Lett. **26**(19), 2913-2916.

Turcotte, D.L. 1999, *Self-organized criticality*, Rep. Prog. Phys. **62**, 1377-1429.

Uritsky, V.M., Klimas, A.J., Vassiliadis, D., Chua, D., and Parks, G. 2002, *Scale-free statistics of spatiotemporal auroral emission as depicted by Polar UVI images: dynamic magnetosphere is an avalanching system*, J. Geophys. Res. **1078**(A12), SMP 7-1, CiteID 1426.

Vassiliadis, D., Anastasiadis, A., Georgoulis, M., and Vlahos L. 1998, *Derivation of solar flare cellular automata models from a subset of the magnetohydrodynamic equations*, Astrophys. J. **509**, L53-L56.

Vlahos, L., Georgoulis, M., Kliuiving, R., and Paschos, P. 1995, *The statistical flare*, Astron. Astrophys. **299**, 897-911.

Voges, W., Atmanspacher, H., and Scheingraber, H. 1987, *Deterministic chaos in accretion systems: Analysis of the X-ray variability of Hercules X-1*, Astrophys. J. **320**, 794-802.

Welinder, P., Pruessner, G., and Christensen, K. 2007, *Multiscaling in the sequence of areas enclosed by coalescing random walkers*, New J. Phys. **9**(5), 149-167.

Willinger W., Govindan, R., Jamin, S., Paxson, V., and Shenker, S. 2002, *Scaling phenomena in the Internet: critically examining criticality*, in *Self-organized complexity in the physical, biological, and social sciences*, Arthur M. Sackler Colloquia, (eds. Turcotte, D., Rundle, J., and Frauenfelder, H.), The National Academy of Sciences: Washington DC, p.2573-2580.

Willis, J.C. and Yule, G.U. 1922, *Some statistics of evolution and geographical distribution in plants and animals, and their significance*, Nature **109**, 177-179.

Wolfram, S. 2002, *A new kind of science*, Wolfram Media, ISBN 1-57955-008-8.

Yoder, M.R., Holliday, J.R., Turcotte, D.L., and Rundle,J.B. 2012, *A geometric frequency-magnitude scaling transition: Measuring b=1.5 for large earthquakes*, Tectonophysics **532-535**, 167-174.

Young, K. and Scargle, J.D. 1996, *The dripping handrail model: Transient chaos in accretion disks*, Astrophys. J. **468**, 617-632.

Yule, G.U. 1925, *A Mathematical Theory of Evolution, based on Conclusions of Dr. J.C. Willis, F.R.S*, Phil. Trans. Royal Soc. of London, Ser B **213** (402-410), 21-87.

Zanette, D.H. 2007, *Multiplicative processes and city sizes*, in "The Dynamics of Complex Urban Systems. An Interdisciplinary Approach", (eds. S. Albeverio, D. Andrey, P. Giordano, and A. Vancheri), Springer: Berlin.

Zipf, G.K. 1949, *Human Behavior and the Principle of Least Effort*, Cambridge MA, Addison-Wesley.

Chapter 3
SOC and Fractal Geometry

R. T. James McAteer

Abstract

Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line -
Benoit Mandelbrot (1982).

When Mandelbrot, the father of modern fractal geometry, made this seemingly obvious statement he was trying to show that we should move out of our comfortable Euclidean space and adopt a fractal approach to geometry. The concepts and mathematical tools of fractal geometry provides insight into natural physical systems that Euclidean tools cannot do. The benefit from applying fractal geometry to studies of Self-Organized Criticality (SOC) are even greater. SOC and fractal geometry share concepts of dynamic n-body interactions, apparent non-predictability, self-similarity, and an approach to global statistics in space and time that make these two areas into naturally paired research techniques. Further, the iterative generation techniques used in both SOC models and in fractals mean they share common features and common problems. This chapter explores the strong historical connections between fractal geometry and SOC from both a mathematical and conceptual understanding, explores modern day interactions between these two topics, and discusses how this is likely to evolve into an even stronger link in the near future.

R. T. James McAteer
New Mexico State University, Department of Astronomy, Las Cruces, NM 88003 e-mail: mcateer@nmsu.edu

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

3.1 From chaos to order (and back)

Probably the most fundamental principle of the entire scientific process is that science and scientific laws are real, are within human understanding, are discoverable, and are universal. Self-Organized Criticality (SOC) is an excellent demonstration of this principal: it is a theoretical construct rooted in reality, discovered by simulations, and universally applicable across a diverse array of topics. SOC is introduced in Chapter 1 of this book; for the purposes of this chapter I would like to summarize it as a forced, non-linear, system that builds up energy via a series of interactions, eventually reaching a critical state at which energy dissipation naturally occurs. These systems tend to exhibit complex patterns in space and time. It is these complex patterns that we use to draw a connection to fractal geometry. It is important to note right from the beginning that although links can be drawn between SOC and fractal geometry, these links are not necessarily causal - i.e., any self-organization will lead to a fractal geometry system and hence SOC is only one such route to fractality.

An interesting historical link between SOC and fractal geometry, and their roles in the scientific process, can be drawn by looking back across human history. Every early form of society and religion have two aspects in common. From the Babylonians, to the Aztecs, to the Indigenous Australians - no matter which early society we look at, we see that they had a similar understanding of the sky (an early geometric cosmology) and of music (an early system of story telling). In his Pulitzer Prize winning book, 'Gödel, Escher, Bach', Hofstadter was one of the first to draw this link between differing forms of human perception and consciousness (Hofstadter, 1970). There are a number of common aspects between these two subjects of music and cosmology, and their link to human perception, which shed some light on the connections between SOC and fractal geometry. First, these early forms of cosmology tended to have an ordered universe appearing out of a disordered void. Early Proto-Indo-European religions often described the formation of the cosmos as the result of a clash between representatives of chaos and order, which of course resulted in a victory for order. It is no surprise that this concept of order appearing *ex-nihilo* has been retained in many modern religions. I discuss how this order-from-chaos applies to fractals and SOC in Section 3.1. Second, both these ideas of early cosmology and music are often considered as attempts by the human brain to force a classification of patterns out of the complexity and noise of everyday life. This classification was well defined by the early Greeks in the form of geometry - a concept that continued to dominate the minds of natural philosophers until the introduction of algebra and calculus in the seventeenth century. I connect fractal geometry to SOC by revisiting the study of geometry in Section 3.2. Third, everyone agrees when they hear beauty in music and see beauty in the sky, but in both cases beauty can often be difficult to quantify. This concept is famously summarized by as,

I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it.
Justice Potter Stewart, Jacobellis v. Ohio, 1964.

I set the scene of *knowing it when we see it* by describing the many and varied attempts at quantifying fractal geometry in relation to SOC in Section 3.3. Fourth, modern cosmology and music show that things tend to be more complicated than at first glance. I discuss the complication of multifractals in Section 3.4, where I attempt to link multiple formulations of multifractality to those of turbulence. Finally, in Section 3.5, I revisit the progression of these studies of SOC and fractal geometry with a look forward at what these two connected branches of science may produce in the future.

3.1.1 The coffee table fractal

In order to probe the connections between fractals and SOC it is useful to first of all study the historical setting of both phrases. In this historical context, fractals and fractal geometry are one of the few mathematical subjects to make the jump from pure abstract equations to the coffee table. In his popular 1982 book 'The Fractal Geometry of Nature' Mandelbrot showed that fractal geometry is universal and natural, occurring everywhere as the output of a system with some forced input and some simple rules (Mandelbrot, 1982). This same rules apply to SOC. The phrase *fractal* was derived slightly earlier by Mandelbrot in his 'Les objets fractals, forme, hasard et dimension' (Mandelbrot, 1975) from the latin *fractus*, meaning fractured or broken. Mandelbrot captured the public's imagination by showing how a simple iterative process can arrived at a complex and beautiful geometry. Just over a decade later in 1987 James Gleick authored another famous best-seller which made the jump from pure mathematics to popular science. In 'Chaos' (Gleick, 1987), Gleick showed the universal nature of chaos theory by demonstrating how a few simple iterative equations rapidly lead seemingly predictable systems into randomness - SOC exhibits a similar behavior of crashes when a critical threshold is met. This was further popularized by the movie 'Jurassic Park', one of the best examples of discussions of the ramifications of chaos theory in all of Hollywood.

The tyrannosaur doesn't obey set patterns or park schedules. The essence of Chaos.

This chaos can be interspersed by periods of order, only to revert back to chaos a few iterations later. It is important to recognize that both these modern subjects of fractals and chaos are connected to that of SOC under the larger and much older umbrella of *dynamics* - in each case we are simply posing the same question as Sir Isaac Newton in 'Principia Mathematica' (Newton, 1687) - How does a system change with time? This gives three potential connections between SOC and fractal geometry.

- Fractal geometry and SOC are both natural and universal.
- Fractal geometry and SOC both have simple inputs, but complex output.
- Fractal geometry and SOC both deal with the study of dynamics.

3.1.2 The n-body problem

Newton made two definitions in Principia that remained unchallenged until Einstein's relativity, and still remain fundamental to studies of SOC and fractal geometry. He stated

> Absolute, true, and mathematical time, of itself, and from its own nature flows equably without regard to anything external. Absolute space, in its own nature, without regard to anything external, remains always similar and immovable.
> Isaac Newton, Principia Mathematica, 1687.

In doing so, he separated space and time. To Newton, space is a stage on which the universe acts, and time shows us how this acting occurs. Models of both SOC and fractals retain this by building a grid (of space) and studying how 'particles' behave in time. Newton famously went on to make numerous substantial contributions in science, including that of calculus. He showed that given sufficiently accurate measurements, he could accurately predict the future evolution of any system. Newton was able to the derive the equations of Kepler (the square of a planet's orbital period varies as the cube of its distance from the Sun) and Galileo (the acceleration of an object is independent of its mass) for any two body system. So Newton's laws could be, and still are, used to predict the Moon's orbit around the Earth, the Earth's orbit around the Sun, or a rocket's trajectory through the atmosphere. However, he also showed the limits of such studies. Try to analyze any three-body system (e.g., the Sun-Earth-Moon), or move to even greater numbers of bodies (e.g., the Sun-Earth-Moon-Rocket), and the equations are unsolvable. We cannot easily predict the position and velocity of more than two interacting bodies. The equations of such an n-body system are non-linear and unstable. Poincaré partially solved this quandary by restating the question in a different form (Poincaré, 1890). He showed that although we may not be able to predict the

position and velocity of each and every body, we can predict the statistical behaviour of such systems. Poincaré also took an early glance down the dark tunnel of chaos theory (although he did not call it 'chaos' as we know it today) by showing that the purely deterministic long term predictability of such an n-body system is impossible. This quick jaunt from Newton to Poincaré gives us three further features of fractal geometry that we can connect to SOC and which we study later in the chapter.

- Fractal geometry and SOC both rely on Newton's definitions of space and time.
- Fractal geometry and SOC both study the global statistics of the n-body problem.
- Fractal geometry and SOC inherently acknowledge that the long-term predictability of n-body systems is impossible.

3.1.3 The butterfly effect

Despite the warnings of Poincaré, many scientists have had long and distinguished careers through the extended study of the n-body problem and non-linear dynamics. These pioneers include vander Pol and Andronov in the theoretical understanding of non-linear oscillators, Birkhoff in the 3-body problem, and Kolmogorov in the the study of turbulence. As is often the case, we can trace the next big breakthrough to the invention of a new piece of technology. In this case the creation of the computer in the 1950s, and the adoption of this by many industries, rapidly led to a new level of understanding of non-linear dynamics. In 1961 Lorenz was modeling the Earth's atmosphere to try to predict the weather, and found that he could not even successfully predict the known behavior in his models. He painstakingly tracked this back to round-off errors in the fourth decimal place of his numerical calculations. These small, seemingly insignificant, errors were creating dramatically different weather patterns in the future behaviour of his models. This is famously coined the 'butterfly effect' where a butterfly flapping its wings in London can eventually cause a storm in Tokyo, and the phrase symbolized the futility of such efforts in weather prediction. However, he also showed that if he plotted the behavior of such a system in an appropriate phase space, the resulting plot would be fractal (the famous strange attractor).

Mandelbrot went on to show that fractal geometry exists everywhere in space and time, from the coast of Norway to the patterns of shells to the shapes of galaxies, and from the behavior of cotton prices to population growth to stellar formation rates. In fact, far from being merely curious and fun (if somewhat unintuitive), fractal geometry is actually the more natural way of describing Nature (Mandelbrot, 1982). Mandelbrot showed that fractal patterns appeared at every possible scale, no matter how much he zoomed into, or zoomed out from, the data. He also described two types of

behavior with clear links to SOC: the Joseph effect, where a seemingly persistent regular series of values is suddenly interrupted; and the Noah effect where massive discontinuities occur on top of a series of noisy data. More than this, Mandelbrot popularized this field, and undoubtably his influence had a Joseph effect on this subject field itself. From a pedagogical perspective, the sort of burst in research activity driven by Mandelbrot's work is a perfect example of an episodic period in revolutionary science as popularized in Kuhn's 'The Structure of Scientific Revolutions' (Kuhn, 1967), which itself is now considered to have had a Joseph effect on the progress of science research. The work of Lorenz and Mandelbrot gives us three further points of emphasis to connect fractal geometry and SOC.

- No matter how seemingly unpredictable any non-linear dynamic system (including SOC) may appear, it can be described by fractal geometry when plotted in the appropriate phase space.
- A snapshot of the spatial distribution within any SOC system may be described by fractal geometry.
- The global dynamic evolution of an SOC system may be described by fractal geometry.

3.1.4 The critical points

The rapid growth and interdisciplinary nature of chaos and fractal geometry in the 1970s and 1980s included many breakthroughs in physics (e.g., Hohenberg, 1971; Swinney & Gollub, 1978; Wolf et al., 1985), biology (e.g., Winfree, 1980), , and probability (e.g., Shaw, 1984). From the perspective of SOC studies, the most important work resulting from this period of rapid discovery was the Bak et al. (1987) paper. They proposed SOC as a mechanism for creating complexity in any natural system. This sandpile model, as described in this book, contains the concepts of self-similarity and scale invariance (described in Section 3.2.2) prevalent throughout studies of fractal geometry. Of course, when studying any natural system by experiments, such scale-invariance must have a lower and upper limit, whereas the previous work in pure mathematics is not restricted to such real world nuances. Further, when combined with Mandelbrot's work, it was clear that the onset of the critical point in these sandpile models resulted in power laws and scale-invariance, which could therefore be well described by fractal geometry in both space and time. However, it is important to realize that this is not the only route to fractal geometry (indeed any self organization naturally leads to power laws and fractal geometry).

The cellular automaton models, as described in this book, showed that the emergence of these critical points is a natural consequence of the system and will always arise no matter what the initial input. Note this non-sensitivity to initial conditions does not disagree with the butterfly effect - the sandpile

models merely showed that critical points will eventually arise, but it does not predict when they will arise. This period of recent development in the fields of SOC and fractal geometry leads us to the final important connections between these two subjects.

- Self similarity and scale invariance are both key aspects of SOC and fractal geometry, and these are statistically linked by power laws.
- Nature provides an upper and lower limit to these power laws that may be absent when presented as pure mathematics.
- Iteratively generated models (including cellular automaton models), are intrinsically fractal. Hence we can connect SOC theories and fractal theories via the study of these models and their application to natural phenomena.

3.2 Fractal Properties

The concept of fractals was conceived in mathematics but has now been so well assimilated into everyday language that its definition is now more commonly non-mathematical. Mathematically, we define a fractal as a set with a fractal dimension which exceeds it topographical dimension, i.e., a set of 0-dimensional points that as a whole is closer to a line, or a 1 dimensional line that is so complex that it is closer to a plane. In this description the topographical dimension is the dimension of the underlying basic shapes that make up the fractal, and the fractal dimension describes the overall complexity, or space-filling nature, of the system. In mathematical terminology, fractals are described by equations which are nowhere differentiable. This definition leads us down the route to non-integer exponents, providing a link to SOC described in Section 3.2.1. In this section we use the similarity dimension in the calculation of the fractal dimension and leave other definitions to Section 3.3.

Fractals are more commonly, and just as accurately, described and comprehended by any lay person without requiring knowledge of abstract mathematics. It is generally defined from a geometrical perspective as a shape each part of which exhibits the same statistical characteristics as the whole, a feature known as self-similarity (Section 3.2.2). In more simple language, a fractal is a shape made up of smaller copies of itself. This creates the scale-free nature of fractals, well described by a power law, and hence the many connections to SOC. It is however important to realize that although all fractals are self similar, not all fractals are the result of an SOC process. This concept of self-similarity leads to the second layman definition of a fractal as a shape which has been generated by an iterative process (Section 3.2.3). By starting with a simple shape, and applying a simple rule, it is trivial to form a fractal. This process leads to the well known coffee-table fractals such as the Sierpinksi carpet and the Koch snowflake. Many cellular automaton

models are generated by the same simple iterative process, hence we expect another close connection between SOC and fractality.

3.2.1 Dimensionality

Classically, we define the dimensionality of an object as the number of independent directions, or the degrees of freedom, at the smallest scale. Hence, we see a point a 0-Dimensional, a line as 1-Dimensional, a drawing or sketch as 2-Dimensional, and a ball or cube as 3-Dimensional. This basis for Euclidean geometry, as defined over 2000 years ago in Euclid's 'The Elements', described everything that existed and was considered the only geometry possible until the discoveries of non-Euclidean space as epitomized by Einstein in the early twentieth century. One way of calculating the dimensionality of any object is to cover the details of the object with smaller and smaller similar objects and count how many of the smaller objects are required to entirely cover the original object at each size. One commonly used method which explains this well is as follows. Take a line of unit length and double it. Clearly it now takes two of the original lines to cover the new line. Now lets take a square of unit size and double it. This makes the square twice as wide and twice as high, so we require four of the original squares to cover it. Try a cube of unit size and double it so it is twice as wide, twice as high, and twice the depth. It will now take eight of the original cubes to cover it. The power of this simple experiment is clear by looking at the results in a table format

Table 3.1: The connection between dimensionally and scaling exponents becomes clear when presented in this Table. N is the number of unit-sized objects required to cover the double-sized object, and d is the dimensionality

Shape	N = 2^d	Dimension
Line	$2 = 2^1$	1
Square	$4 = 2^2$	2
Cube	$8 = 2^3$	3

It is immediately apparent that the Euclidean dimension of an object is equal to the exponent of the scaling factor. This works for any scaling factor including non-integers - try changing the scaling factor to e.g., 3.5, or 1/5 and it will still work. If we move from simple Euclidean shapes to more complex shapes, the analogy seems to fall apart. The most famous example of this is an image of the coastline of Norway. If we use 2-Dimensional boxes to cover an image of this coastline, count them, scale the box, cover the coast with the rescaled boxes, count them, and proceed to rescale the box over and

over we arise at an equation with an exponent of $d = 1.52$. Seemingly the coast is neither 1- or 2- Dimensional. Instead its dimensionality is greater than that of a line, but less than that of a filled-in shape. Interestingly, other countries with smoother coastlines also have non-integer values, albeit much smaller (Coast of Britain $d = 1.25$, Coast of Australia $d = 1.13$). Clearly the fractal dimension of an object is in some way linked to its smoothness (or, in reverse, its spikiness). Fractals give us insight into the complexity of objects that Euclidean dimensions cannot.

We can conceptually connect this to the mathematical definition of differentiation. For example, a curve is differentiable is we can take a gradient at any point along it. However, for a spiky curve there may be locations where the gradient cannot be defined. Now imagine a curve that is spiky everywhere (e.g., the Weierstrass function), one that no matter how many times you zoom in to try to take a derivative, it turns out to be impossible. This curve is, by definition, fractal. These two concepts of dimensionality and differentiability are fundamental to the subject of fractals (Section 3.3.3).

3.2.2 Self-similarity and Scale-Invariance

In nature, we are often confronted with objects with levels of complexity and detail at all scales. Fractal geometry can be used to examine the complexity of these objects across scales. Mathematically, an object is self-similar if it is exactly similar to some part of itself. In nature, few objects are exactly self-similar, but there exists a large family of objects which are statistically self-similar. In the discussion of the the coastlines of countries above we do not expect that the coastline of any part of Norway will look exactly like a miniature form of Norway. Instead it means that, upon zooming into a map of Norway, this zoomed in region will have a number of sharp curves and bends. The statistical distribution of these (i.e., the relative number of sharp curves and smoother bends) will be exactly the same as the country as whole. This form of statistical self-similarity is a typical property of fractals. In comparison, SOC often contains a very specific form of self similarity known as scale invariance, where at any magnification the smaller part of an SOC system is similar to the whole. Hence while a SOC system is always fractal, a fractal system is not always SOC.

3.2.3 Generating Fractals

There are a number of different ways of generating fractals, but by far the most common is that of iterative generation. Here we describe the Cantor set,

the von Koch curve, and how we can solve the inverse problem of describing the rules of generation, given one snapshot in time.

3.2.3.1 The Cantor set and the von Koch curve

The Cantor set is easily formed and contains many interesting facets in a comparison to SOC. It is constructed by starting with a number line containing every possible number between zero and one. Delete the middle third of this number line, but leave behind the end points. This produces two segments, of length 1/3, at each end. Now delete the middle third of each remaining segment, and repeat this operation ad infinitum. The remaining set of numbers at the limit of infinite iteration is the Cantor set. It will be an infinite number of infinitesimally small pieces, with lots of gaps. There are a number of important features of this set from which we can draw parallels to SOC. No matter how many times you zoom into this Cantor set, you will continue to see some structure. Second, it is scale invariant - by zooming in one level and moving to the outside third of the set you will see exactly the same structure as the parent. Third, its fractal dimension is an non-integer, and much higher than 0 (it is actually 0.63). SOC systems are often generated using a similar form of iterative generation, and often exhibit the same precise features of infinitely small structure, scale-invariance, and non-integer exponents.

The von Koch curve is generated by starting with some straight line segment. Remove the middle third and replace it with two sides of an equilateral triangle of the same length as the deleted segment. The resulting curve will have four segments, each length 1/3, for a total length of 4/3. Now repeat this same operation on each segment. The resulting curve from this step will have a length of $(4/3)^2$. Repeat this ad infinitum, and the curve obtained at the infinite iteration is the von Koch curve. The curve will have infinite length $(4/3^{\text{inf}})$ and the arc length between any two segments on the curve is infinite. Its fractal dimension is 1.26. As in the Cantor set, this contains structure at all size scales, is scale-invariant, and has a non-integer fractal dimension. The von Koch curve also shows the problems with trying to use traditional tools in the study of such SOC systems - if we tried to measure the length (or area) of the von Koch curve, we would find it is, rather disturbingly, infinity.

3.2.3.2 Iterative Generation

Iterative generation is a powerful tool for generating a large number of fractals, but is readily explained from a few simple geometric descriptions. The iterative generation of any fractal, including the Cantor set and von Koch curve, is a combination of scaling, reflection, rotation, and translation. Take a simple 2-Dimensional seed shape, where any location of point can be described by its position along an x-axis and a y-axis. Scaling is then the mul-

tiplicative factor in either or both of these directions. Denote these scalings as x' and y', where $x' > 1$ corresponds to an increase in scale and $x' < 1$ corresponds to a reduction in scale. If the scaling is the same in both directions $(x' = y')$, then the resulting fractal will be scale-invariant, otherwise it will be stretched or squashed in each direction. Reflection is described as a sign on the scaling factor, e.g., negative x' means a reflection in the x-axis. Reflection introduces two new components in our transformation matrix. Now, let θ_x be the rotation angle of any horizontal line, and θ_y be the rotation angle of any vertical line. Note that $\theta_x = \theta_y$ provides the conditions to retain scale invariance, otherwise the resulting shape will be stretched or squashed. Finally let Δ_x be a translation along the x-axis and Δ_y be a translation along the y-axis, where by convention up and to the right is positive. These parameters forms a matrix of scaling, reflection, rotation, and translation that prescribe all the freedom necessary to create any fractal shape. Start with the seed shape, apply the transformations in order to generate the second shape, then apply the transformation again to each new segment and repeat. The key link to SOC systems is that this matrix of transformation is usually all that is required to build any self-organized system. For instance, in a simple cellular automaton model, one may simply start off with equi-spaced particles and let them evolve according to some set of rules regarding their position and velocity. Although I described this here using a 2-Dimensional example, this can be extended to any number of dimensions in space and time.

3.2.3.3 The inverse problem

The inverse problem seeks to carry out the reverse procedure to the iterative generation. That is, given the data (position, velocity) of the particles in the system under study, can we calculate the transformation matrix and hence deduce the rules governing the system. As such, the inverse problem is more familiar to most scientists dealing with observational data. In classical mechanics we can analytically solve the inverse problem for a small number of bodies or simple geometry. In simple fractals the inverse problem consists of calculating the matrix which created the fractal geometry. For example the classical snowflake is widely known to have a simple 6-fold rotation symmetry and self-similar scaling. Directly solving for the transformation matrix is quite a powerful technique and can be used to infer the physics behind the generation of more complex fractals in nature, many of which may be the result of SOC, including the the architecture of trees and the classical spiral shape of sea shells and galaxies. Inevitably however, as we include more and more particles, we quickly run into n-body effects. Making an inference of the initial state and deciding on the rules which generated the geometry are almost impossible to obtain. Even if we can obtain such a set of rules, they are almost certainly not unique. In these cases, we can study the statistics of the system and see if it is possible to forward-fit an iteratively generated model to

reproduce these statistics (e.g, in diffusion limited aggregation, particles are modeled to undergo a random Brownian motion and cluster together to form aggregates and simulations can be done on lattices of any desired geometry.)

3.3 The many flavors of fractal dimension

It is instructive to introduce the main types of fractal analyses techniques and dimensions. Often the type of fractal dimension reported is a result of the technique used to obtain it so one must be very careful in cross-comparing any studies of fractal geometry or any comparison of data to models. Here we will introduce two methods of calculating the fractal dimension of both individual systems and a set of these systems, and then define the commonly used phrases of Hurst Exponent, Hölder exponent, and Haussdorff dimension. It is important to understand the similarities and the differences between each method. Further, it is also vital to consider that because of the finite nature of real data, any method can only ever represent an approximation to the actual fractal dimension of Nature.

3.3.1 Similarity Dimension of a set of systems

One method to calculate the similarity dimension of any set of systems is by studying a measure of their 1-Dimensional size compared to their 2-Dimensional size. One such technique relates the perimeter lengths, P, of a large set of structures of different sizes to their areas, A, via the relationship,

$$P(A) \propto A^{D_{PA}/2} , \tag{3.1}$$

where D_{PA} is the perimeter-area similarity dimension of the set of structures. For a simple geometric shape (e.g., a square with side of length l, $A = l^2$, $P = 4l$) $D_{PA} = 1$, whereas for the high perimeter limit (e.g., the same square filled, $A = l^2$, $P = l^2$) $D_{PA} = 2$. Hence D_{PA} is an indication of the complexity of the perimeters of the set of structures and can be calculated from the two-fold slope of a plot of log P versus log A over the entire set of structures. In astrophysics this has been applied to solar granulation in broadband white light images leading to debate over the formation process of these convection features (Roudier & Muller, 1987; Hirzberger et al., 1997; Bovelet & Wiehr, 2001). A closely related study of small scale magnetic fields (Janssen et al., 2003) obtained a similar value of D_{PA} from data and simulations, regardless of the spatial size (i.e., pixel scale) of datasets or the simulation.

The linear size-area technique differs by relating the minimum size of the boxes, L, which contain structures (or clusters of contiguous pixels), to their

area, A, via,

$$L(A) \propto A^{1/D_{LA}} , \tag{3.2}$$

where D_{LA} is the linear size-area similarity dimension of the set of areas. In contrast to the perimeter-area method, for a simple geometric shape $D_{LA} = 2$, whereas for a more serrated shape, $D_{LA} < 2$. A plot of log L versus log A over the entire set of structures will provide an indication of the complexity of the set. This has been applied to sunspot magnetic fields (Lawrence, 1991; Balke et al., 1993) and compared to the predictions of percolation theory. Comparisons of both the perimeter-area and linear size-area methods (Meunier, 1999) to a large set of active regions showed that solar active regions may also be considered as the result of a percolation process at the bottom of the convection zone of the Sun, combined with a diffusion process at the surface. It is worth noting at this stage that such a percolation process is good example of self-organized process that is not fully SOC, but still gives rise to fractal geometry (Section 3.2.2)

3.3.2 Box Counting Dimension

The box counting technique is similar to that described in Section 3.3.1 above, except it can be applied more readily to a single image, (whereas the previous techniques provide the fractal dimension of a set of images). The first step in such a technique is to threshold an image in intensity and thereby create a binary image. This binary image is overlaid with a series of grids consisting of successively larger boxes. The number of boxes, N, containing any part of the binary image scales with box size, ϵ, as

$$N(\epsilon) \propto \epsilon^{-D_{BC}} , \tag{3.3}$$

where D_{BC} is the box counting fractal dimension. The implications of this definition are discussed further in Section 3.3.5, but to complete the comparison to the previous two methods, a simple non-fractal geometric shape will have $D_{BC} = 1$, whereas a non-fractal filled area will have $D_{BC} = 2$. Two extensions of this technique - the Jaenisch and differential box counting methods (Stark et al., 1997) - are designed to overcome the issues with the conversion of the data in the thresholding and the limits of the fractal scaling. The box counting technique can also be extended to provide a multifractal measure (Section 3.4), albeit with significant provisos and problems. The biggest problem in using this technique lies in choosing the threshold value. In a rather simple sense it is possible to threshold the image at different levels of intensity. For standard fractal geometry the resulting calculation of D_{BC} will not vary as a function of intensity thresholds. A change in fractal dimension with changing threshold is often associated with the existence of a structure more complex than a simple fractal.

3.3.3 The Hölder Exponent

Section 3.2.1 introduced the connection of differentiability and fractal geometry. Here we extend this concept by providing a mathematical formalism related to singularities. Singularities are those points where it may not be possible to take a derivative. However, they can still be studied and quantified by analyzing their Hölder Exponent, h. Let $f(x)$ be a function defined on the real number line, and let α be a real positive number. The function f is defined to satisfy the Hölder condition if both C and α exist such that we can find a polynomial P_m of degree m, no greater than the integer part of α, such that for all x in the neighborhood of x_0

$$|f(x) - P_m(x - x_0)| \leq C|x - x_0|^{\alpha} \ . \tag{3.4}$$

Given this, the pointwise Hölder exponent of the function f at x_0 is defined by

$$h(f, x_0) = \sup \{\alpha > 0 | f \in C^{\alpha}(x_0)\} \ , \tag{3.5}$$

that is, the Hölder exponent is the least upper bound such that f belongs to $C^{\alpha}(x_0)$. Consider the following example functions (Seuret & Gilbert, 2000; McAteer et al., 2007),

$$\mathrm{f}_1(x) = \begin{cases} -|x|^3 & \text{for } x < 0 \\ |x|^2 & \text{for } x \geq 0 \end{cases} \ ,$$

$$\mathrm{f}_2(x) = \begin{cases} -|x|^3 \sin(1/x^2) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases} \ .$$

Both functions exhibit sharp changes as one approaches $x = 0$. Note that the derivative of f_1 is continuous at $x = 0$, but that the derivative of f_2 is not defined at $x = 0$. However, for both functions, the Hölder exponent at $x = 0$ is 2 ($h(f, 0) = 2$), since one can write $|f_1(x) - x^2| \leq |x|^2$ and $|f_2(x)| \leq |x|^2$ in this neighborhood. Hence the Hölder exponent permits the characterization of singularities in a series of values, regardless of whether the derivative exists at that point, and provides a powerful technique of calculating the fractal dimension.

3.3.4 The Hurst Exponent

This idea of describing the spikiness (or smoothness) of a function was introduced by Mandelbrot (1982). This work in probability theory originates back to the pioneering work of Robert Brown in 1827, who described the random nature of the motion of pollen grains under his microscope. Mandelbrot defined fractional Gaussian noise as the increments of this Brownian motion phenomena and showed that fractional Gaussian noise signals can be charac-

terized by the Hurst exponent, H. Where c is a constant and $f(x)$ is a fractal process, $g(x) = f(cx)/c^H$ is also a fractal process, moreover it maintains the same statistical distribution. The calculation of the Hurst exponent involves an integration of the signal, therefore when compared to the Hölder exponent, $h = H - 1$. The Hurst exponent describes different types of signals. For Brownian motion, each step along the a series is unrelated to any previous values, the series is just as likely to increase as to decrease and hence we prescribe random white noise with a Hurst exponent of 0.5. Anti-persistent (non-stationary) noise has $0 < H < 0.5$ whereas persistent (stationary) noise has $0.5 < H < 1.0$. Pink noise is described by $H = 1.0$. For a bursty series which tends to show change (i.e., if the series is increasing, the next step is likely to be a decrease and vice versa), the curve covers less distance then a random walk, so $1.0 < H < 1.5$ and the series is described as an *anti-persistent* walk. For a smooth series (i.e., if the series is increasing, the next value will tend to be larger and vice versa), the curve covers more distance than a random walk, so $1.5 < H < 2.0$, and the time series is described as a *persistent* walk.

The Hurst exponent is also intrinsically related to the scaling index of the Fourier power spectral slope of the series. Provided the series is analogous to fractional Gaussian noise and where the Fourier spectral density of the signal scales as a power decay of index -β, then,

$$\beta = 2H - 1 = 2h + 1 \ . \tag{3.6}$$

This provides a powerful method for calculating the Hurst exponent and fractal dimension of any data. One can carry out an FFT, calculate the power spectrum, fit a slope to the energy spectrum and then back out the Hurst exponent. This could then be compared between SOC models and data. However caution is required in assigning strong conclusions to any such study as there are least a dozen methods of calculating the Hurst exponent (Clegg, 2009; Taqqu et al., 1995), and they can differ significantly in their solutions. This is particularly true when calculating the Hurst exponent in data where the intrinsic physical timescale of the system may be much smaller than the resolution of the observations.

3.3.5 Haussdorf Dimension

The description of the Haussdorff dimension provides a more mathematically rigorous and general description of the box counting dimension above (Section 3.3.2). Consider $N(r)$ as the number of balls, radius r, required to cover a n-dimensional signal. As r decreases, $N(r)$ increases such that

$$N(r) \propto \frac{1}{r^D} \ . \tag{3.7}$$

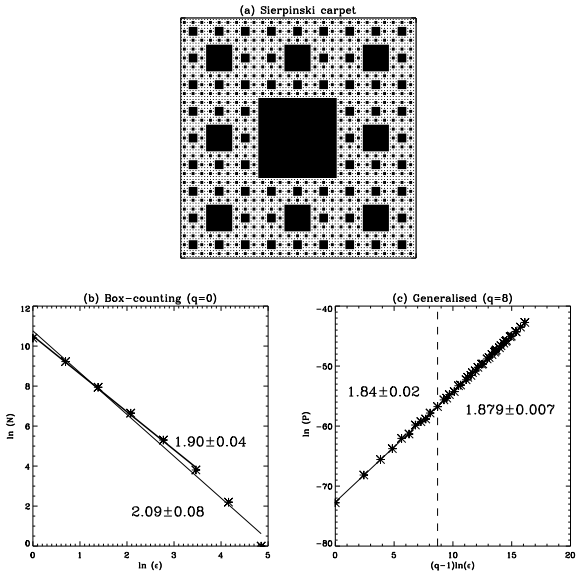


Fig. 3.1: The Sierpinski carpet (top, reproduced from McAteer et al., 2005) is a iteratively generated fractal that highlights the difficulties in calculating the fractals dimension using standard box counting (bottom left) and the possible solutions that can be used to obtain a more accurate determination of its fractal dimension (bottom right)

The Hausdorff-n dimension is then defined as

$$D = -\lim_{r \to 0} \frac{\log N(r)}{\log r} \ . \tag{3.8}$$

The box counting dimension is a version of the Haussdorf dimension implement with boxes rather than circles.

Typically the Hausdorff dimension is the same as the topographical dimension. However, for a fractal the Hausdorff dimension is strictly greater than its topographical dimension (i.e., a seemingly 1-Dimensional fractal curve is so complex, it can be thought of as occupying space closer to 2-Dimensional). For example, a simple straight line on a plane has a Hausdorff dimension of one. But consider a more complex object like the Sierpinski carpet (Figure 3.1) that is a more complex object than a simple line but where it is also clear that it does not fill up the entire plane; the Hausdorff dimension of the Sierpinski carpet is approximately 1.89. Of course for real data it is impossible to approach the limits of $r \to 0$ and so typically Equation 3.8 is described by plotting over a *linear range* of r. In this case the (normally square) image of side length S is covered with grids of boxes with increasing side length ϵ (ϵ belongs to the set $2^n, n = 0, 1, 2, 3...$ such that $\epsilon_{max} = S$)

and $N(\epsilon)$ is the total number of boxes required to cover the image at each box size. The gradient of a straight line fit to the linear section of a plot of log $(N(\epsilon))$ against log (ϵ) corresponds to the fractal dimension. However this simplistic methodology is known to suffer from a poor fit to the log-log plot and so is prone to large errors. This problem is due to the inability of the grid to adequately describe the image at box sizes larger than one pixel, the small number of points in the linear range (a CCD size of $1K \times 1K$ pixels will only have 10 box sizes, $\epsilon_{max} = S = 2^{10}$, therefore the linear range could be as few as $6 - -7$ points), and the often dismissed requirement to rotate the images.

The results of these problems are presented here by studying the Sierpinski carpet in detail. This fractal is created by starting with a uniformly filled square, divided into nine smaller congruent squares, and the interior square removed (i.e., the 8 cells at the edges are uniformly filled with 1 and the centre cell is set to 0). This operation is then repeated on the 8 smaller squares iteratively until no further division can be carried out (the single pixel limit is reached). The resulting fractal has a fractal dimension of $\sim$1.8928 (=log(8)/log(3)). A simplistic box-counting algorithm produces the plot in Figure 3.1b. Although the shape is fractal, the log(N)–log(ϵ) plot appears concave - a multifractal signature (Section 3.4). Over the entire range of scales, the fractal dimension is calculated as 2.09 ± 0.08 . Judging by eye, there appears to be a break at a box size of 32 pixels, and the gradient of the curve over the box size range of 1–32 pixels gives a much better estimation (1.90 ± 0.04) of the true fractal dimension.

There are a few steps which we can implement to circumnavigate these problems (McAteer et al., 2005; Lawrence et al., 1996). At each box size, $N(\epsilon)$ is calculated as the *minimum* number of boxes required to cover the image. This is achieved by translating the grid to every possible origin (e.g., a grid of boxes of side $\epsilon = 3$ will have 9 different possible origins). This should be repeated for every orientation of the image, but the obvious difficulties in precisely rotating a pixel array by any non-90 degree amount make this rotation difficult once the data is in this form (this should ideally be carried out in experiment by rotating the camera and taking snapshots at each rotation). In addition we can remove the 2^n box size dependence (which was really only introduced to decrease computation time) and remove the need to have square boxes (i.e., for a rectangular box of sides a and b, the effective box size is $\epsilon_{eff} = \sqrt{ab}$).

This dramatically increases the number of data points at small box size (note the large number of data points in Figure 3.1c), hence we can introduce an quantitative indication of the linear section of the fit. For L effective box sizes we carry out a linear fit to the log-log plot in a sliding window of size l $(L/2 \leq l \leq L)$. At each position of the window, the fit is accepted if the spread in the fit in small; a window with a significant non-linear contribution will produce a large spread. This is displayed in Figure 3.1c where a fit over the entire range of box sizes produces a fractal dimension of 1.84$\pm$0.02 (an

underestimation of the true fractal dimension). A fit over a smaller section of the curve produces a fractal dimension of 1.879±0.007 (a much better estimation of the true fractal dimension), hence this section contains less of the non-linear range. Note I have also adopted a higher order moment in calculating the value Figure 3.1c), an approach borrowed from the study of multifractals.

3.4 Multifractals

It is entirely possible that any given signal will contain many different singularities strengths (Hölder exponents) at many different positions. For a signal with infinite resolution (or at least resolution smaller than any physical event) it would be possible to measure the Hölder exponent of each event and create a number density histogram to reflect the different numbers of each exponent. However when singularities are not isolated, as is generally the case for real data, one may calculate the Hausdorff dimension of the different singularities in the data. In this case, the Hausdorff dimension provides the necessary *statistical mechanics* approach to describing the number density of each Hölder exponent and performs the role of quantifying the degree with which a particular Hölder exponent is evident in a signal. For example, if the Hausdorff dimension of a particular Hölder exponent was 1, this would mean that the singularity of this type is evident everywhere in the time series. This naturally leads onto the topic of multifractals - shapes with a fractal geometry more complex than that of simple fractals. Although fractals are useful for quantifying the space-filling nature of any n-dimensional system, it has been found that many systems in nature are a convolution of different fractal processes. A multifractal exists when the measure itself is self-similar (Evertsz & Mandelbrot, 1992), and is usually described as a concave curve known as the singularity spectrum, $f(\alpha)$. If we are to make a comparison between SOC and measures of fractal geometry we must ensure we address this issue of multiple fractal processes within one signal. In this section I will expand on this discussion, with particular emphasis on the generalized dimensions, turbulence, and on the wavelet transform.

3.4.1 From monofractals to multifractals

Section 3.3 showed that the fractal dimension of any object can be thought of as the self-similarity of a signal across all scale sizes. For any 2-Dimensional image this can be described as the scaling index of any length to area measure,

$$A \propto l^{\alpha} \ , \tag{3.9}$$

where α is the singularity strength. However, a multifractal system will contain a spectrum of singularity strengths of different powers,

$$A \propto l^{f(\alpha)} \ , \tag{3.10}$$

and takes account of the measure at each point in space. In Section 3.3.2 a fractal measure was obtained by thresholding in intensity and creating a binary image, hence the value of the measure in each pixel was lost. The route to producing multifractal measures starts with retaining these pixel intensities in the form of a measure distribution. In general, the measure distribution is characterized by

$$\psi(q,\tau) = E\sum_{i=1}^{N} P_i^q \epsilon^{-\tau} \ , \tag{3.11}$$

where q and τ can be any real numbers, and E is the mean value of the measure with N components. In this form, ψ is the coupled τ -moment of the size ϵ, and q-moment of the measure P. The three main multifractal indices commonly used to represent a non-uniform measure are then the:

- generalized dimensions Grassberger & Procaccia (1983), $D_q = \tau/(q-1)$;
- singularity strength, $\alpha = d\tau/dq$;
- legendre transformed $f(\alpha) = q\alpha - \tau$ / .

When applied via a traditional box-counting approach, it is useful to define the partition function,

$$Z_q(\epsilon) = \sum_{i=1}^{N} P_i^q(\epsilon), \tag{3.12}$$

such that $\tau(q) = lim_{\epsilon\to 0} log(Z)/log(\epsilon)$, and any of the three representations above can be calculated. The q moment plays the role of increasing the relative importance of the more intense parts of the measure as q is increased. In this way it acts as a microscope to investigate the different contributions made to the image at higher values of the measure.

3.4.2 Generalized Dimensions

We can arrive at the multifractal spectrum by generalizing the previous concept of the Hausdorff dimension as a member of an infinite series of q^{th} order dimensions. The classical generalized q^{th} order fractal dimension, D_q, of any image is given (e.g., Grassberger & Procaccia, 1983; Hilborn, 2000) as,

$$D_q \ = \frac{1}{q-1} \ lim_{\epsilon\to 0} \frac{ln \ \Sigma_{i=1}^{n} P_i(\epsilon)^q}{ln \ \epsilon} \ , \tag{3.13}$$

where $q = -\infty, ..., +\infty$. For a shape consisting of m pixels, P_i is the probability of finding a pixel in the i^{th} box (with a total of n boxes) and is given by,

$$P_i(\epsilon) \ = \ \frac{m_i(\epsilon)}{m} \ , \tag{3.14}$$

for a total of m points in the image: hence P_i is a indication of the measure in the i^{th} box.

A measure is defined as the magnitude of the values within an area, e.g., for an image it could be the sum of the pixel values within the box under study (compare this with the simple box counting approach where the actual pixel values are thrown out when the binary image is constructed at some chosen threshold.) Each measure is then normalized to ensure that the sum of all the measures is unity. The exponent q is used to extract the information at each value of the measure - when q is positive it magnifies the larger measures, dwarfing the smaller ones; when q is negative it inverts the measures, thus enhancing the smaller measures and dwarfing the larger ones. The normalized measure is given as,

$$\hat{P}_\epsilon(i) = \frac{P_\epsilon(i)^q}{\Sigma_{(i=1)}^{N(\epsilon)} P_\epsilon(i)^q} = \frac{P_\epsilon(i)^q}{z_\epsilon(q)} \ , \tag{3.15}$$

where $z_\epsilon(q)$ is once again the partition function. The fractal strength α is then

$$\alpha(q) = lim_{\epsilon \to 0} \Sigma_{i=1}^{N\epsilon} \hat{P}_\epsilon(i) log_\epsilon P_\epsilon(i) \ . \tag{3.16}$$

Equation 3.16 provides the local fractal dimension for $P_\epsilon(i)$, and is a measure of the strength of the fractal dimension at that scale q. The full singularity spectrum $f(\alpha)$ is then given by repeating this over a range of values of q,

$$f(q) = lim_{\epsilon \to 0} \Sigma_{i=1}^{N\epsilon} \hat{P}_\epsilon(i) log_\epsilon \hat{P}_\epsilon(i) \ , \tag{3.17}$$

$$f(\alpha) = q\alpha - lim_{\epsilon \to 0} \Sigma_{i=1}^{N\epsilon} \hat{P}_\epsilon(i) log_\epsilon z_\epsilon(q) \ , \tag{3.18}$$

$$f(\alpha) = q\alpha - \tau \ , \tag{3.19}$$

where we now define $\tau = lim_{\epsilon \to 0} log_\epsilon E(z_\epsilon(q))$. So $f(\alpha)$ is a weighted sum of the log of the normalized measure to the base ϵ.

3.4.3 Connecting forms of multifractality

Although there are multiple descriptions of multifractality, they are readily connected by a few simple equations (McAteer et al., 2010). Here we formulate and connect the Haussdorff dimension, Hölder exponent, equations of thermodynamics, and structure functions.

3.4.3.1 The Hausdorff dimension and Hölder exponent

As we see below in section 3.4.5, one approach to measure the multifractal dimension is to calculate the Hausdorff dimension, D, of each Hölder exponent, h in a signal. The resulting $D(h)$ spectrum as calculated at each moment, q, is directly related to the $f(\alpha)$ spectrum as,

- $f(\alpha) = D(h_q)$,
- $h_q = \alpha - E_{dim}$,

where E_{dim} is the Euclidean dimension ($E_{dim} = 2$ for an image).

3.4.3.2 The link to thermodynamics

The multifractal spectrum is often generated from the Legendre transform of the scaling exponents of a signal using an analogy to the laws of thermodynamics. The $f(\alpha)$ singularity spectrum was originally introduced to provide a statistical description of the multifractal spectrum in a manner related to well known thermodynamic quantities. In this analogy with thermodynamics, f is 'entropy', α is 'internal energy', and these are related via a Legendre transform to τ ('free energy') and q ('temperature'),

$$\alpha = \frac{d\tau}{dq} \ , \tag{3.20}$$

$$f(\alpha) = q\alpha - \tau(q) \ , \tag{3.21}$$

and this can be further related to the spectrum of generalized dimensions,

$$D_q = \tau(q)/(q-1) \ . \tag{3.22}$$

As it is generally simpler to calculate the D_q spectrum, this relationship can be used to produce the singularity spectrum (e.g., Vlahos et al., 1995). In this case, the signal is covered by boxes (of increasing size l) and it found that the qth moment of the measure (or probability of occurrence), P, in each box, i, scales as,

$$D_q(q-1) = \tau(q) = \lim_{l \to 0} \frac{\log \sum_i P_i^q(l)}{\log(l)} \ . \tag{3.23}$$

By further analogy with thermodynamics, the summation is known as the partition function where, as already describe earlier, positive q will accentuate the large values, and negative q will survive the small values, of the measure. The calculated D_q is then Legendre transformed to calculate the multifractal spectrum. However, this method suffers from known problems of non-linear scaling and discontinuities in D_q. A second approach circumvents this problem by calculating the singularity in each box i, $\alpha_i = \log \ P_i(l)/\log \ l$. The histogram $N(l)$ then varies as,

$$N(l) \sim l^{-f(\alpha)} \ , \tag{3.24}$$

and $f(\alpha)$ can be directly calculated. Unfortunately this method is shown to be inaccurate due to slow convergence (Chhabra & Jensen, 1989).

3.4.3.3 Structure Functions

Another method of calculating the multifractal spectrum is based on the structure function (Parisi & Frisch, 1985; Abramenko et al., 2002). This method consists of calculating the statistical moments of the field increments $S_q(r)$, as a function of separation, r in order to determine the scaling exponents, ζ_q,

$$S_q(r) \sim r^{\zeta_q} \ , \tag{3.25}$$

which are also directly related to the multifractal spectrum as (Muzy et al. 1993),

$$D(h) = qh - \zeta_q + 1 \ , \tag{3.26}$$

$$h = \frac{d\zeta}{dq} \ . \tag{3.27}$$

In equation 3.27, it is the deviation of h away from a single value that signifies multifractality.

3.4.4 The Devils staircase

It is instructive to look at a typical multifractal spectrum and discuss a few salient features. The $D(h)$ multifractal spectrum of a well-known multifractal, the devil's staircase (also known as the Cantor function), is displayed in Figure 3.2, and the basic properties are as follows.

- The Hölder exponent (abscissa), as described in Section 3.3.3 is the singularity strength. The Hausdorff dimension (ordinate), as described in Section 3.3.5 reflects the space-filling degree of each Hölder exponent.
- The concave shape is typical of many multifractal spectra. The left leg describes the fewer, larger amplitude, events, which correspond on large positive q, whereas the right leg reflects the more common, smaller, singularities described by large negative q.
- The Hölder exponent with the largest Hausdorff dimension is the most common singularity strength and can be approximated to the Hurst exponent of the entire series, ($h = 0.63$, $H = 1.63$). The Hausdorff dimension of this Hölder exponent, $D(h) = 0.61$, is the Capacity Dimension and quantifies the space-filling degree of this singularity: $D(h) = 1$ describes a signal

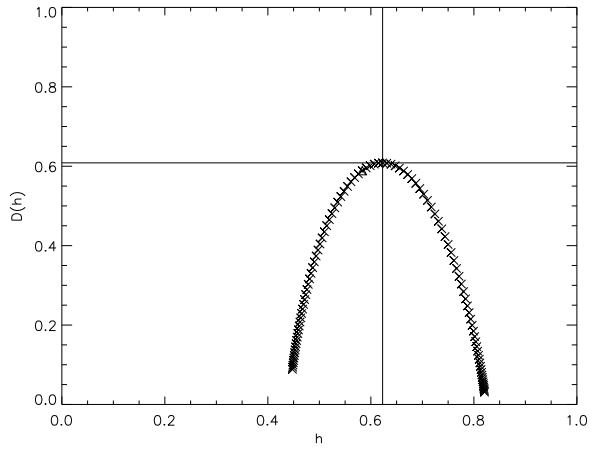


Fig. 3.2: The multifractal spectrum of the devil's staircase, represented as the Hausdorff dimension, D, of each Hölder exponent, h, exhibits the usual concave structure.

that is singular almost everywhere; $D(h) < 1$ describes a signal with rarer singularities.
- The spectrum exists over a range of Hölder exponents [$h_{min} = 0.47$, $h_{max} = 0.84$]. In the limit of monofractality, the spectrum would collapse to single point, only one Hölder exponent is necessary, and the range tends to zero.
- The degree of symmetry of the curve about $D(h)$ reflects the degree of inhomogeneity between rarer large amplitude and more common, small amplitude, singularities.

3.4.5 The Wavelet Transform Modulus Maxima

The box counting technique (Section 3.3.2) has been applied successfully in multifractal studies (Lawrence et al., 1993; Cadavid et al., 1994) and a change in fractal dimension with scale (i.e., changing threshold) is often associated with the existence of a multifractal structure. However, it is important to note the difference between a true multifractal and the often observed sudden change of fractal dimension at large scales which may result from algorithm imperfections and/or data issues. A modification of this method to try to account for this is to randomly sample the image and renormalize the calculations (Cadavid et al., 1994). Other modifications include the Ĺ‡fuzzy ballĂ“ method (Alber & Peinke, 1997) which involves shaking the boundary of the box, and averaging over these local measures to obtain a more accu-

rate result of the true measure. However, there are know drawbacks to each of these methods. Box counting methods are known to be threshold dependent and structure functions are not defined for negative values of q. As such both methods are somewhat limited in their application to real data sets (Conlon et al., 2008, 2009; Georgoulis, 2005).

The continuous wavelet transform provided a new and natural method to overcome these issues and perform a complete multifractal analysis (Parisi & Frisch, 1985). The wavelet transform modulus maxima (WTMM) method replaces the boxes from the traditional box counting method with wavelets that act as fuzzy boxes, are defined for both finite and discrete domains, and are better suited for real data sets. Application of the WTMM method to 1-Dimensional time series have provided insight into a wide variety of problems, e.g., fully-developed turbulence, financial markets, meteorology, physiology, and DNA sequences (Muzy et al., 1991, 1993; Arneodo et al., 1995, 2002). The WTMM methods were generalized to 2 dimensions for multifractal analysis of rough surfaces and associated physical systems (Decoster et al., 2000) and to 3-dimensional turbulence dissipation data (Kestener & Arneodo, 2004).

The WTMM was formalized to overcome the algorithmic difficulties associated with Legendre transform methods and takes advantage of the large body of work carried out on the wavelet transform. The wavelet transform decomposes a N-dimensional signal $f(t)$ in both time and scale by the convolution of the signal with a set of dilated and translated wavelets,

$$WT(b,a) = \frac{1}{a^{N/2}} \int_{-\infty}^{\infty} \psi^* \left[\frac{x-b}{a} \right] f(t) dt \ . \tag{3.28}$$

In this equation, ψ^* is the complex conjugate of the mother wavelet, a plays the role of adjusting its width, and b shifts the wavelet along the signal. The mother wavelet is chosen to be well localized in both space and time, but furthermore if ψ has n vanishing moments, it can be shown that all polynomials up to order $n-1$ will be convolved to zero, thereby unmasking the local Hölder exponent by removing the $n-1$ polynomial trend. In the series of derivatives of a Gaussian,

$$\psi^n(t) = \frac{d^{(n)}}{dt^n} e^{\frac{-t^2}{2}} \ , \tag{3.29}$$

the first n moments are vanishing. The wavelet power spectrum can be averaged over time to produce a global wavelet spectrum analogous to the Fourier energy spectrum. This requires assigning a Fourier frequency to each wavelet scale and normalizing both the Fourier and global wavelet spectrum (Torrence & Compo, 1998) (e.g., for the $n = 2$, Mexican hat, wavelet, this increases the minimum observed scale from $2 \times dt$ to $7.95 \times dt$). The further benefit of using a wavelet decomposition comes from the creation of the modulus maxima tree. At each scale, localized maxima in the modulus of the wavelet transform are identified. These are then connected across scales to

form maxima lines, essentially ridges identifying maxima across scale. There is always at least one WTMM line pointing towards any singularity (Mallat, 1998), t_0, and this scales as,

$$WT(t_0, a) \sim a^{h(t_0)+\frac{1}{2}} \ . \tag{3.30}$$

When singularities are well separated in time, this can be used to calculate the local Hölder exponent for each singularity (Struzik, 1998; Struzik, 2000). However, when singularities are not isolated, a global partitioning has to be introduced. Let t_{mm} be all positions of local maxima of $|WT(t, a_0)|$ at some scale a_0, then define the partition function as the summation of the modulus of the wavelet coefficients to some power, q,

$$Z(q, a) = \sum_{mm} |WT(t_{mm}, a)|^q \ . \tag{3.31}$$

As before, the moment, q, plays the role of increasing the weighting of contribution of the largest wavelet transform coefficients to the summation when $q > 0$, and increasing the weighting of the smallest wavelet transform coefficients to the summation when $q < 0$. The approach of only including the modulus maxima removes very small values of the wavelet transform, which would otherwise cause divergence at $q < 0$. This also dramatically reduces the computational time, and incorporates the branch-like multiplicative structure of the wavelet transform into the partition function. Although the multifractal spectra may be calculated from this partition function, it is generally preferred to used the canonical approach (Chhabra & Jensen, 1989; Muzy et al., 1991) where,

$$u_{mm_i}(q, a) = \frac{|(WT(t_{mm_i}, a)|^q}{\sum_{mm_i} |WT(t_{mm_i}, a)|^q} \ , \tag{3.32}$$

is the effective Boltzmann weighting of each modulus maxima risen to q at each scale. From this new measure,

$$h(q) + \frac{1}{2} = \lim_{a \to 0} \frac{1}{\log a} \sum_{mm_i} u_{mm_i}(q, a) \ \log \ |(WT(t_{mm_i}, a)| \ , \tag{3.33}$$

$$D(h(q)) = \lim_{a \to 0} \frac{1}{\log a} \sum_{mm_i} u_{mm_i}(q, a) \ \log \ (u_{mm_i}(q, a)) \ . \tag{3.34}$$

Equations 3.33 and 3.34 can be solved by a linear regression of the summation against the scale values ($log\ a$), and hence the singularity spectrum, $D(h)$, can be directly calculated at each q. The log-log plots of Equation 3.33 and 3.34 are generally linear, with more scatter about the linear trend at negative q. The major problem with the WTMM approach is in tracking the ridges correctly. Maxima in the wavelet may exist which do not correspond

to singularities; these ridges will not exist down to the very lowest scales. Tracking rules which are too relaxed will include these maxima, thereby contaminating the Boltzmann weightings in Equations 3.33 and 3.34. On the other hand, tracking rules which are too strict may not be able to follow the ridge to large enough scales; again this will affect the Boltzmann weightings in Equations 3.33 and 3.34. Another recent advance in this field uses the properties of the wavelet transform in order to overcome the limitations of the box counting method (Conlon et al., 2009; Kestener et al., 2010) by pre-selecting those modulus maxima corresponding to background noise and removing them before forming the multifractal spectrum.

3.5 Future directions

Since Mandelbrot first introduced the fractal dimension, the idea of quantitatively describing the complexity of a system (or an image of that system) has been applied in many areas of science. From a purely theoretical viewpoint, the fractal dimension is an indication of the self similarity of a shape across multiple size scales. We can conceptually consider multifractals as a conglomeration of fractals occupying the same point (in space, time, or both space and time). From a mathematical perspective, the links between SOC and fractal geometry are numerous - many SOC systems can be described by fractal geometry when described in appropriate phase space, any snapshot of an SOC system is likely to be fractal, and the long term dynamic evolution of an SOC system is likely to be fractal. The features that provide these connections are self-similarity (and scale invariance), power laws, the upper and lower limits to the scale-free range and the iterative methods used to generate both SOC models and fractals. It seems clear that fractal geometry provides one route to test our modern SOC models against real data. However, there are caveats - there is more than one fractal dimension and the techniques used to calculate fractal geometry are not all generally applicable. Recent advances in generating accurate and robust multifractal measures may provide a new tool to link models to data. Specifically, the incorporation of multiscale techniques (e.g., wavelets) may assist with the problems of spatial and temporal resolution. They also provide connections to theories in the areas of turbulence and magnetohydrodynamics that assist us in inferring physical parameters from unitless indices. The final argument for increasing research in this area of connecting fractal geometry and SOC may be that of a new data-driven approach to science. Our data volumes (from both models and experiments) and computational ability are increasing beyond the capabilities of an individual scientist and even an individual research field. The interdisciplinary ability to quantify features and classify images is set to become a vital tool in the scientists kit. Perhaps the links between fractal geometry and SOC can lead the way in these efforts.

References

Abramenko, V.I., Yurchyshyn, V.B., Wang, H., Spirock, T.J., Goode, P.R. Scaling Behavior of Structure Functions of the Longitudinal Magnetic Field in Active Regions on the Sun, ApJ, 577, 487 (2002)

Alber, A., Peinke, J., An Improved Multifractal Box-Counting Algorithm, Virtual Phase Transitions, and Negative Dimensions (1997)

Arneodo, A., Bacry, E., Graves, P.V., Muzy, J.F., Characterizing long-range correlations in DNA sequences from wavelet analysis, Phys. Rev. Lett. 74, 3293 (1995)

Arneodo, A., Audit, B., Decoster, N., Muzy, J.F., Vaillant, C. Wavelet based multifractal formalism: Application to DNA sequences, satellite images of the cloud structure and stock market data in The Science of Disasters: Climate Disruptions, Heart Attacks, and Market Crashes (A. Bunde, J. Kropp, H. J. Schellnhuber, eds.), pp. 26-102. Springer Verlag, Berlin. (2002)

Balke, A. C., Schrijver, C. J., Zwaan, C., Tarbell, T. D. Percolation theory and the geometry of photospheric magnetic flux concentrations, Sol. Phys., 143, 215 (1993)

Bak, P., Tang, C., Wiesenfeld, K., Self-organized criticality: An explanation of the 1/f noise. Physical Review Letters 59 (4), 381, (1987)

Bovelet, B., Wiehr, E. A New Algorithm for Pattern Recognition and its Application to Granulation and Limb Faculae, Sol. Phys., 201, 13 (2001)

Cadavid, A.C., Lawrence, J.K., Ruzmaikin, A.A., Kayleng-Knight, A. 1994, ApJ, 429, 391

Conlon, P. A., Gallagher, P. T., McAteer, R. T. J., et al., Sol Phys, 248, 297 (2008)

Conlon, P.A., Kestener, P., McAteer, R.T.J., Gallagher, P.T., Fennel L., Khalil, A. Arneodo, AA (2009)

Chhabra, A., Jensen, R., Direct determination of the f(alpha) singularity spectrum, Physical Review Letter 62 (12), 13271330 (1989)

Clegg, Richard, A practical guide to measuring the Hurst parameter, ISSN 1473-804x online (2009)

Decoster, N., Roux, S.G., Arneodo, A. A wavelet-based method for multifractal image analysis. II. Applications to synthetic multifractal rough surfaces Eur. Phys. J. B 15, 739-764. (2000)

Evertsz, C.J.G., Mandelbrot, B.B., Multifractal measures in: Peitgen, H.-O., Jué rgens, H., Saupe, D. (Eds.), Chaos and Fractals. Springer-Verlag, New York, (1992)

Georgoulis, G.,Turbulence In The Solar Atmosphere: Manifestations And Diagnostics Via Solar Image Processing, 228, 5 (2005)

Gleick, James, Chaos, USA (1987)

Grassberger, P., Procaccia, I., Measuring the Strangeness of Strange Attractors. Physica D: Nonlinear Phenomena 9 (12), 189, (1983)

Hirzberger, J., V‡zquez, M., Bonet, A., Hanslmeier, A., Sobotka, M.. Time Series of Solar Granulation Images. I. Differences between Small and Large Granules in Quiet Regions, ApJ, 480, 406 (1997)

Hofstadter, Douglas, Gödel, Escher, Bach: An Eternal Golden Braid, Basic Books, USA (1970)

Hohenberg, Pierre, Dynamical theory of critical phenomena, in E. G. D. Cohen (Ed.) Statistical mechanics at the turn of the decade, Dekker, New York (1971)

Hilborn, R., Chaos and Nonlinear Dynamics: An introduction for Scientists and Engineers Oxford University Press, UK (2000)

Janssen, K., Všgler, A., Kneer, F., On the fractal dimension of small-scale magnetic structures in the Sun, AA, 409, 1127 (2003)

Kestener, P., Arneodo, A. Generalizing the wavelet-based multifractal formalism to random vector fields: Application to three-dimensional turbulence velocity and vorticity data Phys. Rev. Lett. 93, 044501 (2004)

Kestener, P. Conlon, P.A. Khalil, A., Fennell, L., McAteer, RTJ.,Gallagher, PT., Arneodo, A. Characterising Complexity in Compound Systems: Segmentation in Wavelet-Space, ApJ, 717, 995 (2010)

Kuhn, Thomas, The structure of scientific revolutions, USA (1967)

Lawrence, J. K. Diffusion of magnetic flux elements on a fractal geometry, Sol. Phys., 135, 249 (1991)

Lawrence, J.K., Ruzmaikin, A.A., Cadavid, A.C. 1993, ApJ, 417, 805

Lawrence, J.K., Cadavid, A.C., Ruzmaikin, A.A. On the Multifractal Distribution of Solar Magnetic Fields, ApJ, 465, 425 (1996)

Mandelbrot, Benoit B, The Fractal Geometry of Nature, San Francisco (1982).

Mandelbrot, Benoit B, Les objets fractals, forme, hasard et dimension, Paris, Flammarion (1975)

Mallat, S. G. A Wavelet Tour of Signal Processing, San Diego Academic, (1998)

McAteer, R.T.J., Gallagher, P.T, Ireland, J., Statistics of Active Region Complexity: A Large-Scale Fractal Dimension Survey, ApJ, 631, 628 (2005)

McAteer, R.T.J., Young, C.A., Ireland, J., Gallagher, P.T, The Bursty Nature of Solar Flare X-Ray Emission, ApJ, 662, 691 (2007)

McAteer, R.T.J., Gallagher, P.T., Conlon., P.A., Turbulence, complexity, and solar flares, Advances in Space Research, 45, 9, 1067 (2010)

Meunier, N. Fractal Analysis of Michelson Doppler Imager Magnetograms: A Contribution to the Study of the Formation of Solar Active Regions, ApJ, 515, 801 (1999)

Muzy, J.F., Bacry, E., Arneodo, A. Wavelets and multifractal formalism for singular signals: Application to turbulence data Phys. Rev. Lett. 67, 3515-3518, (1991)

Muzy, J.F., Bacry, E., Arneodo, A., Multifractal formalism for fractal signals: the structure fonction approach versus the wavelet transform modulus maxima method Phys. Rev. E 47, 875 (1993).

Newton, Isaac, Principia Mathematica, UK (1687)

Parisi, G., Frisch, U., On the singularity structure of fully developed turbulence in Turbulence and Predictability in Geophysical Fluid Dynamics (eds. M. Ghil, R. Benzi and G. Parisi). Proceedings of the International School of Physics ŇEnrico FermiÓ 8487, North-Holland. (1985)

PoincarŽ, Jules Henri, Sur le problŹme des trois corps et les Žquations de la dynamique. Divergence des sŽries de M. Lindstedt. Acta Mathematica, 13, 1270 (1890)

Roudier, Th., Muller, R. Structure of the solar granulation Sol. Phys., 107, 11 (1987)

Seuret, S., Gilbert, A. Proc. 13th ITC Specialist Seminar on Internet Traffic Measurement and Modelling (Paris: Int. Teletraffic Congr.) (2000), No. 14

Shaw, R. The Dripping Faucet As a Model Chaotic System, Science Frontier Express Series, Aerial Press, (1984)

Stark, B., Adams, M., Hathaway, D. H., Hagyard, M. J. Evaluation of Two Fractal Methods for Magnetogram Image Analysis Sol. Phys., 174, 297 (1997)

Struzik, Z. R. Removing Divergences in the Negative Moments of the Multi-Fractal Partition Function with the Wavelet Transformation, CWI Rep. INS-R9803, Amsterdam: Cent. Wiskunde Inf., (1998)

Struzik, Z. R., Siebes, A. P. J. M.. Outlier Detection and Localisation with Wavelet Based Multifractal Formalism, CWI Rep. INS-R0008, Amsterdam: Cent. Wiskunde Inf. (2000)

Swinney, H.L. and Gollub, J.P. Hydrodynamic Instabilities and the Transition to Turbulence, Physics Today, 1978

Taqqu, M., Teverovsky, V., Willinger, W. Estimators for long-range dependence: an empirical study, Fractals, 3, 785 (1995)

Torrence, C., Compo, G. P., Bull. Am. Meteorol. Soc., 79, 61 (1998)

Wolf, A, Swift, J. B, Swinney, H. L, Vastano, J. A. Determining lyapunov exponents from a time series. Physica D: Nonlinear Phenomena 16, 285317 (1985)
Winfree, A.T, The geometry of biological time, springer, new york (1980)
Vlahos, L, Georgoulis, M., Kluiving, R., Paschos, P., The statistical flare, AA, 289, 897 (1995)

Chapter 4
Percolation Models of Self-Organized Critical Phenomena

Alexander V. Milovanov

Abstract In this chapter of the e-book "Self-Organized Criticality Systems" we summarize some theoretical approaches to self-organized criticality (SOC) phenomena that involve percolation as an essential key ingredient. Scaling arguments, random walk models, linear-response theory, and fractional kinetic equations of the diffusion and relaxation type are presented on an equal footing with theoretical approaches of greater sophistication, such as the formalism of discrete Anderson nonlinear Schrödinger equation, Hamiltonian pseudochaos, conformal maps, and fractional derivative equations of the nonlinear Schrödinger and Ginzburg-Landau type. Several physical consequences are described which are relevant to transport processes in complex systems. It is shown that a state of self-organized criticality may be unstable against a bursting ("fishbone") mode when certain conditions are met. Finally we discuss SOC-associated phenomena, such as: self-organized turbulence in the Earth's magnetotail (in terms of the "Sakura" model), phase transitions in SOC systems, mixed SOC-coherent behavior, and periodic and auto-oscillatory patterns of behavior. Applications of the above pertain to phenomena of magnetospheric substorm, market crashes, and the global climate change and are also discussed in some detail. Finally we address the frontiers in the field in association with the emerging projects in fusion research and space exploration.

Alexander V. Milovanov
Associazione EURATOM-ENEA sulla Fusione, Centro Ricerche Frascati, Via Enrico Fermi 45, I-00044 Frascati (Rome), Italy, EU e-mail: alexander.milovanov@enea.it

Also at:
Department of Space Plasma Physics, IKI-Space Research Institute, Russian Academy of Sciences, 84/32 Profsoyuznaya street, 117997 Moscow, Russian Federation e-mail: amilovan@iki.rssi.ru

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

Table 4.1: Abbreviations used in this chapter

Abbreviation	Expansion
AE	Alfvén eigenmode
AGU	American Geophysical Union
AMPTE	Active Magnetospheric Particle Tracer Explorers
ac	alternating-current
AO	Alexander-Orbach (conjecture)
BTW	Bak-Tang-Wiesenfeld
CA	Cellular Automation
CTRW	Continuous Time Random Walks
ECRH	Electron Cyclotron Resonance Heating
ENSO	El Niño / La Niña-Southern Oscillation
EPM	Energetic Particle Mode
DANSE	Discrete Anderson Nonlinear Schrödinger equation
DEMO	DEMOnstration Power Plant
DIII-D	DIII-D tokamak
DP	Directed Percolation
DPRW	Dynamic Polarization Random Walk
IKI	Space Research Institute, Moscow, Russia
ISSI	International Space Science Institute, Bern, Switzerland
ITER	International Thermonuclear Experimental Reactor
FAST	Fusion Advanced Studies Torus
FDE	Fractional Diffusion equation
FGLE	Fractional Ginzburg-Landau equation
FNLSE	Fractional Nonlinear Schrödinger equation
FTU	Frascati Tokamak Upgrade
KAM	Kolmogorov-Arnold-Moser
KWW	Kohlrausch-Williams-Watts (relaxation function)
LH	lower hybrid (oscillation)
L-H	low-high (transition)
MHD	Magnetohydrodynamics
MW	Megawatt
NLSE	Nonlinear Schrödinger equation
ROY	IKI-led “Swarm” project
SOC	Self-Organized Criticality

4.1 The Percolation Problem

The standard theory of percolation (Broadbent and Hammerslay 1957) began with an attempt to make statistical predictions about the possibility for a fluid to filter through a random medium, predictions that could be applied to a variety of physics problems, such as epidemic processes with and without immunization, the underground spread of pollution, and electrical discharges in thunderstorms. The phenomenon is characterized by a finite threshold, to be associated with a critical concentration of fractures, pores, or other sort of conducting channels in the medium, below which the spread is limited to a finite domain of ambient space, and is unlimited otherwise. The percola-

tion problem is relevant for a number of transport problems with threshold behavior as for instance Anderson localization (Anderson 1958) and hopping conduction in amorphous solids (Shklovskii and Efros 1984). The percolation transition is perhaps the simplest phase transition-like phenomenon, with the macroscopic connectedness thought as a spontaneously occurring property, and the concentration of conducting elements as the control parameter (i.e., the analog thermodynamical temperature) (Isichenko 1992).

4.1.1 Site and Bond Percolation

Given a periodic lattice, embedded in a d-dimensional Euclidean space, one can choose between two alternative formulations of the percolation problem: site and bond. The differences between site and bond percolation are actually very subtle and are manifest in a typically lower threshold for the bond problem. There also exists a hybrid, site-bond percolation due to Heermann and Stauffer (1981). In site percolation one assumes that the lattice sites are occupied at random with the probability p (and hence with the probability $1-p$ are empty). A connected cluster is defined as a collection of all occupied sites that can communicate via the nearest-neighbor rule. In bond percolation, one thinks of clusters of connected conducting bonds instead. In this formulation all sites are initially occupied and bonds are occupied randomly with the probability p. Statistically, the p value decides on how big the connected clusters could be for the given topology of the lattice. The key point of the theory is the existence of a critical value, the percolation threshold p_c ($0 < p_c \leq 1$), above which the connected clusters span the entire lattice with the probability 1 (for $d \geq 2$, the threshold value is typically a fraction between 0 and 1; for $d = 1$, $p_c = 1$). The critical probability being smaller than 1 implies that the infinite clusters do not fill the ambient space yet. For $p < p_c$, the percolation dies away exponentially. The threshold value is non-universal: it depends on the type of the percolation problem (site, bond, or hybrid); details of the lattice (cubic, diamond, triangle, etc.); as well as the ambient dimensionality $d \geq 1$. The typical realizations of site and bond clusters on a square lattice are illustrated in Fig. 4.1. It is worth remarking that any point belonging to the infinite percolation cluster can be connected to the infinitely remote point via a connected escape path which lies everywhere on the cluster. (For comprehensive reviews on percolation see, e.g., Stauffer 1979; Stauffer and Aharony 1992; Isichenko 1992; Nakayama et al. 1994; Havlin and ben-Avraham 2002.)

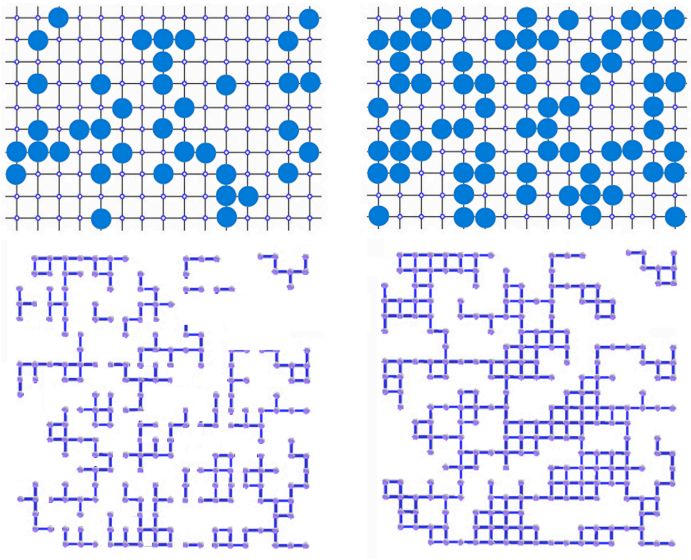


Fig. 4.1: Site vs. bond percolation. Top: Site percolation problem, with circles representing occupied sites on a lattice. Left: A lattice system below the percolation threshold. Right: The same lattice system at the threshold of percolation, with a noticeably denser concentration of the occupied sites. Bottom: A similar picture for bond percolation. Connected (conducting) bonds are shown in blue color. Left: An insulating state of the lattice below the percolation threshold. Right: Random distribution of conducting bonds at the threshold of conducting dc electricity.

4.1.2 Percolation Critical Exponents β, ν, and μ

Likewise to traditional critical phenomena, characterized by a scale-free statistics of the spontaneously occurring quantities, the geometry of connected clusters in vicinity of the percolation threshold is self-similar (fractal) (e.g., Stauffer 1979; Feder 1988). As $p \to p_c$, the percolation correlation (i.e., pair connectedness) length diverges as $\xi \propto |p - p_c|^{-\nu}$. For $p > p_c$, the probability to belong to the infinite cluster is $P_\infty(p) \propto (p - p_c)^\beta \propto \xi^{-\beta/\nu}$, whereas the dc conductivity behaves as $\sigma_{\rm dc} \propto (p - p_c)^\mu \propto \xi^{-\mu/\nu}$. The critical exponents β, ν, and μ are universal in that they do not depend on the type of the percolation problem, nor on details of the lattice. They do depend on the ambient dimensionality d, however, and their numerical values are known in all $d \geq 1$. The "mass" of a connected cluster scales with its size as $M(\xi) \propto \xi^d P_\infty(p) \propto \xi^{d-\beta/\nu}$, leading to a nontrivial Hausdorff dimension

$d_f = d - \beta/\nu$. The latter expression is sometimes said to be "hyperuniversal" as it holds in any $d \geq 1$.

4.1.3 Random Walks on Percolating Clusters

The problem of diffusion on fractals (de Gennes 1976; Straley 1980; Gefen et al. 1983) has stirred considerable attention in the literature, especially, in terms of the random walk approach. If the random walker (an unbiased "ant") is put on a connected cluster at percolation, then the distance it travels after t time steps behaves as (Gefen et al. 1983; reviewed in ben-Avraham and Havlin 2000; Havlin and ben-Avraham 2002)

$$\left\langle \mathbf{r}^2(t) \right\rangle \propto t^{2/(2+\theta)}. \tag{4.1}$$

The exponent θ is given by $\theta = (\mu - \beta)/\nu$. Note that the dependence here is no longer proportional to the time t, in contrast to uniform spaces. Thus, diffusion is *anomalous*. The exponent θ describes topological characteristics of the fractal (such as connectivity, etc.) It shows, moreover, remarkable invariance properties under smooth (diffeomorphic) maps of fractals (contrary to the Hausdorff dimension d_f, see Milovanov 1997; Zelenyi and Milovanov 2004). In the literature, this exponent is referred to as both the connectivity exponent and the index of anomalous diffusion. This ambiguity merely reflects that the diffusion is anomalous because fractals possess anomalous connectedness features as voids are present at all scales. When $\theta \to 0$, normal (Fickian) diffusion is introduced.

In a basic theory of percolation (Stauffer 1979; Stauffer and Aharony 1992) it is shown that $\mu > \beta$ for connected clusters, implying that $\theta > 0$. One sees that the mean-square displacement in Eq. (4.1) grows slower-than-linear with time. This slowing down of the transport occurs as a result of multiple trappings and delays of the diffusing particles in cycles, bottlenecks, and deadends of the fractal object on which the random motions concentrate. Note that the scaling law above holds as a single-cluster rule (the "ant" cannot jump between the clusters). Averaging over all clusters at percolation replaces Eq. (4.1) by

$$\left\langle \mathbf{r}^2(t) \right\rangle \propto t^{(2-\beta/\nu)/(2+\theta)} \tag{4.2}$$

for $t \ll \xi^{2+\theta}$. Equation (4.2) has implications for the ac conductivity at "anomalous" frequency scales, $\omega \gg \xi^{-(2+\theta)}$, for which the charge carriers move only on the fractal (Gefen et al. 1983; Milovanov and Rasmussen 2001). The various aspects of anomalous diffusion in fractal systems are summarized in the reviews (Bouchaud and Georges 1990; ben-Avraham and Havlin 2000; Havlin and ben-Avraham 2002; Zelenyi and Milovanov 2004).

4.1.4 The Spectral Fractal Dimension

A hybrid parameter $d_s = 2d_f/(2+\theta)$ is often referred to as the spectral, or fracton, dimension. It is so called because it represents the density of states for vibrational excitations in fractal networks termed fractons (Alexander and Orbach 1982; Rammal and Toulouse 1983; reviewed in Nakayama et al. 1994). It also appears in the probability of the random walker to return to the origin ($\propto t^{-d_s/2}$) (e.g., O'Shaughnessy and Procaccia 1985). The key difference between the Hausdorff and spectral fractal dimensions lies in the fact that d_f is a purely structural characteristic of the fractal, whereas d_s mirrors the dynamical properties, such as wave excitation, diffusion, etc., which enter via the connectivity exponent. Note that, because $\theta \geq 0$ for percolation, the spectral fractal dimension is not larger than its Hausdorff counterpart, i.e., $d_s \leq d_f$. The value of d_s can conveniently be considered as an effective fractional number of the degrees of freedom in fractal geometry, as it naturally substitutes the integer (embedding) dimension in respective diffusion (Gefen et al. 1983; O'Shaughnessy and Procaccia 1985) and wave-propagation (Alexander and Orbach 1982; Orbach 1989; Nakayama et al. 1994) problems on fractals.

4.1.5 The Alexander-Orbach Conjecture

In the past years there has been much excitement about the Alexander-Orbach (AO) conjecture that the spectral fractal dimension is exactly 4/3 for percolation clusters in any ambient dimension d greater than 1 (Alexander and Orbach 1982; reviewed in Nakayama et al. 1994; Havlin and ben-Avraham 2002). This conjecture is important as it relates the structural characteristics of the fractal, contained in d_f, to the dynamical characteristics, contained in θ. For $d \geq 6$, the AO conjecture was proven by Coniglio (1982) as a percolation problem on a Cayley tree (Bethe lattice). A Cayley tree is a graph without loops where each node contains the same number of branches (called the coordination number). In many ways, owing to its intrinsic hierarchical structure, a Cayley tree behaves as an infinite-dimensional space (its volume grows exponentially fast with the scale, in contrast to power-law growth for physical lattices, fractal or not, see Schroeder 1991). Not surprisingly, the percolation problem on a Cayley tree is regarded as a suitable model for mean-field percolation. For $d < 6$, the mean-field approach is invalidated as loops become important at all scales, thus impeding reduction to the trees. A great deal of effort has been invested to prove or disprove the AO conjecture in the lower embedding dimensions $d < 6$. It is now clear that in these dimensions the AO conjecture is not exact, nor does it generalize to all statistical fractals as for instance to the backbones of percolation clusters (Stanley and Coniglio 1984; Havlin and ben-Avraham 2002). Even so, the

"true" values found for the spectral fractal dimension at criticality continue to be numerically surprisingly close to the original AO result $d_s = 4/3$ for all $d \geq 2$ thus sustaining the conjecture (ben-Avraham and Havlin 2000).

4.1.6 Percolation Problem on the Riemann Sphere

It is both interesting and instructive to demonstrate how the spectral fractal dimension may be obtained for threshold percolation on a plane ($d = 2$). The main idea here (Milovanov 1997) is to extend the plane on which the percolation is considered by adding the point at infinity to it, then perform a stereographic projection of the infinite percolation cluster on the Riemann sphere (see Fig. 4.2). As a result the percolation problem will be compactified, since the point at infinity is mapped to the north pole.[1]

Observe that the percolation cluster spanning the plane implies that its stereographic image covers part of the surface of the sphere, the north pole included. Without loss of generality, we may assume that the south pole at which the sphere touches the plane belongs to the cluster. When considered on the Riemann sphere, a percolating escape path to infinity originating from the south pole will be a simple arc connecting the two poles, south to north. The key step is to notice that a solid angle at the base of this arc has a lower bound as posed by connectedness. Indeed this angle cannot be smaller than the angle at the base of the half meridian. The latter angle is immediately seen to be equal to π. Clearly, the percolation cluster itself is based on a solid angle not smaller than this. It is convenient to think of the number d_s of the degrees of freedom as corresponding to an orthogonal basis of d_s vectors (Milovanov 1997), which span a fractional solid angle $\Omega_{d_s} = d_s \pi^{d_s/2} / \Gamma(d_s/2+1)$. Here, Γ denotes Euler's gamma function. Partial cases of this expression are, $\Omega_2 = 2\pi$ for $d_s = 2$ and $\Omega_3 = 4\pi$ for $d_s = 3$. Thus, we expect that, for connected clusters, $\Omega_{d_s} \geq \pi$, from which a lower bound on d_s may be deduced by defining $\Omega_{d_s} = \pi$. We associate this lower bound with the threshold of macroscopic connectedness (threshold of percolation). In equating Ω_{d_s} to π we used that the stereographic projection being a conformal map is angle and circle preserving. Putting all the various pieces together, we have (Milovanov 1997; reviewed in Zelenyi and Milovanov 2004)

$$d_s \frac{\pi^{d_s/2}}{\Gamma(d_s/2+1)} = 2\frac{\pi^{d_s/2}}{\Gamma(d_s/2)} = \pi. \tag{4.3}$$

[1] Essentually the same ideas are used in the complex analysis to extend the complex plane, a procedure known as the one-point compactification. Remember that the stereographic projection is a conformal map, which is angle and circle preserving. Here, we apply the technique of one-point compactification to the percolation problem, aiming to generate an object which might be described as a compact percolation cluster (Milovanov 1997).

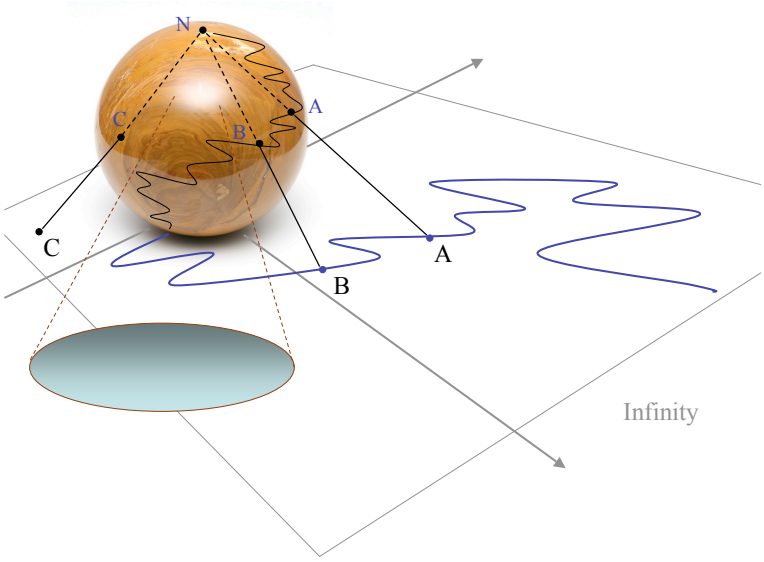


Fig. 4.2: Stereographic projection of a percolating cluster on the Riemann sphere. The north pole, N, represents the point at infinity. When considered on the Riemann sphere a percolating escape path to infinity (blue line, with points A and B on it) originating from the south pole at which the sphere touches the plane is a simple arc connecting the two poles, south to north. A solid angle at the base of this arc has a lower bound, π, as dictated by connectedness. Then the mapping being conformal implies that at the threshold of percolation $\Omega_{d_s} = \pi$, leading to $d_s = 1.327 \pm 0.001$.

Numerical solution shows that $d_s = 1.327 \pm 0.001$, remarkably close to, although slightly smaller than, 4/3. Rigorously speaking, this result disproves the AO conjecture in $d = 2$. Despite being this subtle, the observed deviation from 4/3 is important as it helps to avoid the secular terms problem when applying a renormalization-group technique near the percolation point (Nakayama et al. 1994; Havlin and ben-Avraham 2002). Note that the solution to Eq. (4.3) has a remarkable meaning. It defines the fractional dimensionality of a ball-like space seen from its center under the solid angle π. It is the dimensionality of this space, $d_s \approx 1.327$, which permits percolation in terms of a connected escape path to infinity. One sees that the percolation problem is essentially a topological problem. It decides on macroscopic connectedness of random systems in terms of the number of the coupled degrees of freedom. As such, the percolation problem has important implications for the dynamics of complex systems, and self-organized critical systems as particular case, as it will be demonstrated shortly.

4.1.7 Summary

Summarizing, the percolation problem presents a non-trivial problem with scale-free behavior. It is related with phase transition-like phenomena as well as the fundamental topology (via the connectedness issues). The percolation clusters provide a particularly clear example of statistical fractals in the limit $\xi \to \infty$. The percolation indices β, ν, and μ are known in any ambient dimension $d \geq 1$ and do not depend on details of the lattice nor on the type of the percolation problem (universality). Topologically, the percolation problem decides on the existence of a connected escape path to infinity. Focusing on the percolation problem in $d = 2$ ambient dimensions, it is convenient to extend the Euclidean plane on which the percolation is considered, by adding the point at infinity to it (i.e., the one-point compactification, used in the complex analysis). Then the percolation problem will be "compactified" in the sense that the point at infinity will belong to the infinite cluster. Accordingly, the extended Euclidean plane with the "compact" percolation cluster on it can conformally (by using a stereographic projection) be mapped onto the Riemann sphere. This compactification procedure when applied to percolation presents a significant appeal: Not only does it permit to keep the point at infinity "at hand" by identifying it with the sphere's north pole − the stereographic projection being a conformal map makes it possible to analytically obtain the value of the spectral fractal dimension at criticality beyond the mean-field approaches. This value, which is approximately equal to ≈ 1.327, is a transcendental number and it has a remarkable meaning (Milovanov 1997): It defines the fractional dimensionality of a ball-like space seen from its center under the solid angle π. This result offers some insight into the mathematical origin of the spectral fractal dimension. In addition to basic science, the percolation problem is of practical importance as it offers a platform for the description of transport properties of disordered (random) media. In particular, diffusion and electrical conduction problems on percolation clusters have been widely studied and discussed in the literature (e.g., Nakayama et al. 1994; Havlin and ben-Avraham 2002; Zelenyi and Milovanov 2004; references therein).

4.2 The SOC Hypothesis

The challenge to understand fractals (Mandelbrot 1982) and the ubiquitous $1/f$ "noise" led Bak, Tang, and Wiesenfeld (BTW, 1987, 1988) to introduce the concept of self-organized criticality, or SOC. The claim was that irreversible dynamics of systems with many coupled degrees of freedom ("complex" systems) would naturally generate self-organization into a critical state without fine tuning of any external or control parameter(s). By analogy with traditional critical phenomena it was argued that in vicinity of the critical

state there is universal behavior, robust with respect to variations of parameters and with respect to randomness, and that the system can be characterized by power-laws and a set of critical exponents. An impressive list of publications[2] have been produced in the attempt to prove or disprove the SOC hypothesis for the various systems. The phenomenon was demonstrated on a number of automated lattice models, or "sandpiles," displaying avalanche dynamics and scale invariance (Bak et al. 1987, 1988; Tang and Bak 1988; Zhang 1989; Kadanoff et al. 1989). The various aspects of self-organized criticality dynamics have been reviewed by Bak (1996); Jensen (1998); Turcotte (1999); Charbonneau et al. (2001); and Aschwanden (2011).

To qualify as SOC, the system must be open, be coupled with the exterior, and involve many interacting degrees of freedom. In addition, its dynamics must be thresholded and nonlinear, and the driving, or energy injection, rate must be very slow (infinitesimal). An important advance of SOC is the realization that fractals appear naturally through a self-organization process and that the corresponding critical state is an attractor for the dynamics. In many ways the notion of SOC can be thought of as belonging to the nascent "science of complex systems" which addresses the commonalities between apparently dissimilar natural, technological, and socio-economic phenomena, e.g., market crashes and climate disruptions (e.g., Bunde et al. 2002; Albeverio et al. 2006). Despite its promising performance, the SOC hypothesis is a subject of strong debate in the literature, and many issues related to it remain controversial or in demand for further investigation.

4.2.1 SOC vs. Percolation

Before we proceed with the main topics of this chapter, we would like to address the SOC hypothesis against the percolation problem discussed above. Indeed SOC shares with percolation the implications of threshold behavior and spatial self-similarity. An essential difference is that percolation is a purely geometrical model, whereas SOC involves, in addition, the temporal counterpart of the fractal, the $1/f$ noise (e.g., Montroll and Shlesinger 1982). In many ways SOC is a *spatio-temporal* phenomenon where both spatial and temporal self-similarities are coupled and long-ranged.

Another important aspect is that in percolation and other traditional critical phenomena, control parameters must be fine tuned to obtain criticality (thus the name "control"). In SOC phenomena, control parameters make part of the dynamical system instead: their values are defined dynamically as the system self-adjusts to accommodate the changing exterior conditions. It is in this sense that a SOC system is said to "automatically" (without a fine tuning

[2] According to ISI's Web of Science the number of papers citing the seminal work by Bak, Tang, and Wiesenfeld (1987) as by the fall of 2012 is at three and a half thousand.

of parameters) reach the critical state.[3] More so, we remark that there exists a "self-organized" formulation of some standard percolation processes such as the spread of deceases or forest fires. It was argued that their dynamics could be formulated so that they mimic SOC phenomena (Grassberger and Zhang 1996). The characteristic feature here is that singularities at p_c emerge not in distribution of order parameters but of control parameters, making these phenomena look like SOC. This formulation is advantageous, as it leads to efficient numerical algorithms, allowing for a precise determination of the critical behavior, as for instance in models of self-organized critical directed percolation, with time interpreted as the preferred dimension (Maslov and Zhang 1996).

4.2.2 The Guiding Mechanisms

An important issue concerns the mechanisms that "guide" a system to criticality. These mechanisms are of two types. One type is associated with the application of an extremal principle that the dynamics should obey in order to satisfy the microscopic equations of motion. Examples of this type are the invasion percolation (Wilkinson and Willemsen 1983) and Bak-Sneppen models (Bak and Sneppen 1993; reviewed in Paczuski et al. 1996). In invasion percolation[4] – introduced in physics by Wilkinson and Willemsen (1983) – the dynamics proceed along a path of least resistance under the action of capillary forces. Under the condition that the flow rate is infinitesimal the system finds its critical points that are stable and self-organized. The second type is associated with the operation of a feedback mechanism between system's dynamical parameters (Kadanoff 1991) as for instance a feedback of the order parameter on the control parameter(s) as discussed by Sornette (1992a) and Gil and Sornette (1996). The Bak, Tang, and Wiesenfeld's (BTW) sandpile is a prominent example of this type. In sandpiles the unstable sand slides off to decrease the slope and reinstall stability, thus providing a feedback of the particle loss process on the dynamical state of the pile. Generally, self-organized criticality can occur in every system with a negative feedback mechanism. Introducing a parallel with traditional thermodynamics, a thermostat is designed in such a way that the same temperature is maintained through fluctuations, so that the system is self-controlling.[5] The self-organized nature of the criticality stems from the fact that the spatial correlations being long-ranged act as attracting the nonlinear feedback dynamics (Sornette 1992a). We should stress that a feedback plays a very important role in the phenom-

[3] Often one says that a SOC system possesses no tunable control parameters, but that's all about the wording.

[4] To be distinguished from ordinary percolation discussed above.

[5] Here, the notion of a "temperature" is thought as analog control parameter in SOC phenomena.

ena of SOC as it ensures a steady state where the system is marginally stable against a disturbance (Kadanoff 1991). Following Sornette (1992a), we also note that using the idea of feedback it is possible to convert the standard critical phenomena into self-organized critical dynamics, thereby extending considerably the span of models exhibiting SOC. We illustrate this in section 4.7.7, where a "self-organized" localization-delocalization transition on a separatrix system of nonlinear Schrödinger equation with randomness is considered.

4.3 Going With the Random Walks: DPRW Model

The percolation problem when account is taken for a dynamical feedback mechanism offers a suitable platform to build toy-models of self-organized critical phenomena. Early attempts in this direction refer to the "dilution-by-hungry-ants" and "thermal-fuse" models (with and without a healing) (Sornette 1992a). In what follows, we discuss a model (Milovanov 2010, 2011), dubbed dynamic polarization random walk (DPRW) model, which combines the implication of a feedback mechanism with the idea of random walks on a fractal cluster at percolation. The model is formulated as a transport problem for electrically charged particles of different kinds.[6] The advent of random walks in place of automated lattice redistribution rules makes it possible to calculate the frequency-dependent complex susceptibility of the dynamical system at SOC together with the memory (response) function and in the end to obtain the SOC critical exponents in terms of three percolation critical indices β, μ, and ν. This approach paves the way for an analytical theory of SOC starting from the microscopic dynamical properties. One by-product of the random-walk model is a demonstration (Milovanov 2010, 2011) that the relaxation of a supercritical system to SOC is of Mittag-Leffler type (Mittag-Leffler 1905; reviewed in Metzler and Klafter 2000) (similar to the Cole-Cole behavior in glassy systems and polymers: see Coffey 2004). The Mittag-Leffler relaxation implies that the behavior is multi-scale with a broad distribution of durations of relaxation events consistently with a description in terms of the fractional relaxation equation (e.g., Metzler and Klafter 2000; Sokolov et al. 2002) and at odds with a single-exponential relaxation dynamics of the Debye type (Coffey 2004 for an overview; references therein).

[6] This electrical context of the model is non-crucial and can be relaxed. See section 4.5.2.

4.3.1 Description of the Model

We consider a hypercubic d-dimensional ($d \geq 1$) lattice confined between two opposite $(d-1)$-dimensional hyperplanes, which form a parallel-plate "capacitor" as shown in Fig. 4.3. The plate on the right-hand-side is electrically grounded. Free charges are built by external forces on the capacitor's left plate. When a unit free charge is added to the capacitor the lattice responds by burning a unit "polarization" charge, which is an occupied site added at random to the lattice. When a unit free charge is removed from the capacitor a randomly chosen occupied site is converted into a "hole" site (missing occupied site). A hole will be deleted from the system (converted into empty site) if/when the corresponding free charge has reached (or been moved to) the ground level. There is a limit, $Q_{\max}$, on the amount of the free charges the capacitor can store, and this is defined as $Q_{\max} = ep_cN$, where e is the elementary charge ($e = -1$), p_c is the percolation threshold, and N is the total number of sites across the lattice. Thus, the ability of the capacitor to store electric charges is limited to the occurrence of the infinite cluster at the percolation point. If, at any time, the above limit is exceeded, a double amount[7] of the free charges in excess of $Q_{\max}$ will be removed from the capacitor and will be distributed between the sites of the infinite cluster with equal probability. The implication is that the capacitor leaks electric charges above the percolation point. This property reflects the onset of the dc conduction at the threshold of percolation.

When a hole appears on the infinite cluster it causes an activation event with the following consequence: One of the nearest-neighbor occupied sites, which is a random choice, will deliver its charge content to the hole. The hole which has just received the polarization charge becomes an ordinary occupied site, while the donor site becomes a hole. The newborn hole, in its turn, will cause a further activation event at the location where it has occurred, thus sustaining/triggering a chain reaction of redistribution of polarization charges. The chain reaction continues until the hole reaches the grounded plate where it is absorbed (converted into an empty site). When a hole appears on a finite cluster it causes a chain reaction of activation events in much a similar way as on the infinite cluster, but with an extra condition regarding the termination of the activation process: The chain reaction stops if (i), likewise to the infinite cluster case, the hole reaches the grounded plate where it is converted into empty site, or if (ii) there are no more activities going on on the infinite cluster. In the latter case the finite cluster freezes in a "glassy" state with the quiescent holes in it until either a new hole appears on the infinite cluster or one or more occupied site are added to the lattice by external forces.

[7] This mimics non-zero inductance in the conduction process.

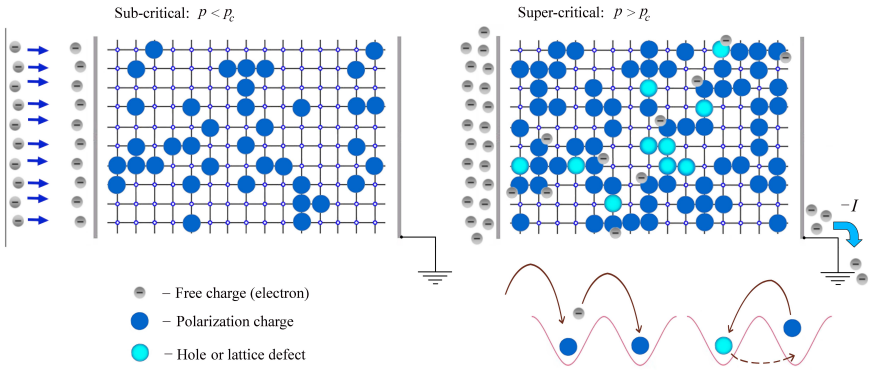


Fig. 4.3: Dynamic polarization random walk (DPRW) model. Grey, blue, and azure particles show respectively the free charges, polarization charges, and holes. Left: System below the percolation point. Free charges are built by external forces on the capacitor's left plate. Right: System slightly above the percolation point, with an illustration of hopping activities on the lattice. Adapted from Milovanov (2011).

4.3.2 Random-Walk Hopping Process

Essentially, the holes interchange their position with the nearest-neighbor occupied sites, and it appears reasonable to model this process as interchange hopping process (e.g., Dyre and Schrøder 2000). We shall assume, following de Gennes (1976), that there is a characteristic microscopic hopping time, which is taken to be unity, but more general hopping models can be obtained by introducing a distribution of waiting times between consecutive steps of the hopping motion (continuous time random walks, or CTRW's) (Montroll and Weiss 1965; Schneider and Wyss 1989). With the above assumption that the site acting as donor is a random choice the transport model is defined as a random-walk hopping model. Similarly to the hole case, the free charges are assumed to behave as unbiased random walkers after their re-injection on the infinite cluster. They will hop at a constant rate between the nearest-neighbor occupied sites in random direction on the cluster on which they are initially placed until they reach the electrically grounded plate where they sink into the ground level of the circuit (see Fig. 4.3). The holes act as conducting sites for the motion of the free charges. The charged plate acts as a perfectly reflecting boundary (as opposed to the absorbing boundary at the electrically grounded plate). Hops to empty sites are forbidden. The latter condition limits the random walks to fractal geometry of the threshold percolation.

4.3.3 Dynamical Geometry of Threshold Percolation

Overall, one can see that the system responds by chain reactions of random-walk hopping processes when it becomes slightly supercritical and it is quiescent otherwise. Excess free charges dissipating at the grounded plate provide a feedback mechanism by which the system returns to the percolation point. There will be a slowly (as compared to hopping motions) evolving dynamical geometry of the threshold percolation resulting from the competition between the adding of occupied sites to the lattice and the charge-releasing chain reactions. Based on the quantitative analysis below, we identify this state as a SOC state. This general picture based on the idea of a dynamic polarization response with random-walk hopping of the charge carriers has been called dynamic polarization random walk (DPRW) model (Milovanov 2010, 2011).

In the DPRW SOC model, chain reactions of the hopping motion acquire the role of "avalanches" in the traditional sandpiles. In the present analysis, we are interested in obtaining the critical exponents of the DPRW model by means of analytical theory. Numerical simulation of the DPRW dynamics is under way for comparison with the analytical predictions. By the time this chapter is being written, the characteristic signatures of multi-scale conductivity response of the dynamical system at criticality have been confirmed in the computer simulation model. In Fig. 4.4, we illustrate the existence of relaxation events of various sizes due to hole hopping on a 10×10 square lattice with random distribution of the conducting nodes and the probability of site occupancy such as to mimic the percolation threshold and the conjectured SOC activities.

4.4 Linear-Response Theory

4.4.1 Dynamics and Orderings

Starting from an empty lattice (no potential difference between the plates), by randomly adding occupied sites to it, one builds the fractal geometry of the random, or uncorrelated, percolation, characterized by three percolation critical exponents β, ν, and μ (connected clusters have fractal dimensionality $d_f = d - \beta/\nu$) (Stauffer 1979; Isichenko 1992; Nakayama et al. 1994; Havlin and ben-Avraham 2002). Remark that the infinite percolation cluster, in the true sense of the wording, exists only in the thermodynamic limit when the lattice itself is infinite. This limit arises because of the need to model the system-sized conducting clusters in terms of fractal geometry. In the absence of holes this percolation geometry is static (polarization charges can only move by exchanging their position with a hole) but when the holes appear on the lattice they cause local rearrangements in the distribution of the conduct-

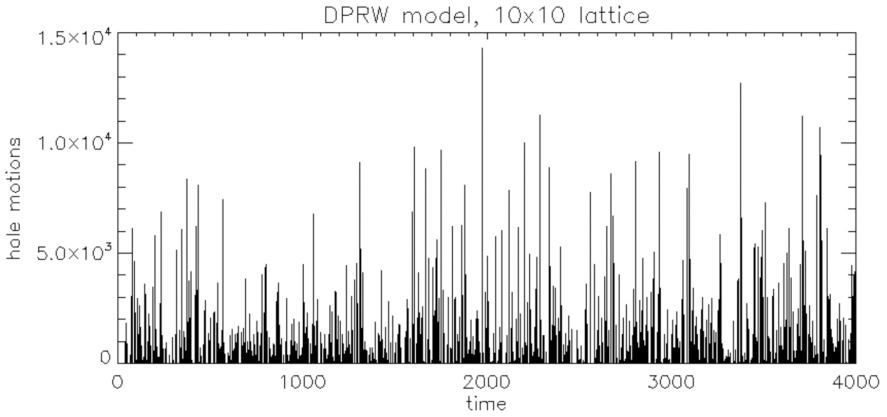


Fig. 4.4: Sizes of chain reactions due to hole hopping on a 10×10 square lattice. Each spike corresponds to a chain reaction of hole hopping on the system-scale conducting cluster, with the height proportional to the number of hops to absorption at the grounded plate. This simple numerical realization illustrates the existence of relaxation events of various sizes consistently with the implication of SOC. Adapted from Milovanov (2011).

ing sites. As a consequence, the conducting clusters on which the transport processes concentrate change their shape and their position in the real space. In the analysis of this section we shall require that the average number density of the holes be very small compared to the average number density of the polarization charges. The implication is that the system remains near the percolation point despite the slow evolution of the conducting clusters. Note that the lattice rules are such as to preserve the properties of the random percolation. In fact, no correlations are introduced in the distribution of the conducting sites at any step of the lattice update.

4.4.2 Frequency-Dependent Conductivity and Diffusion Coefficients

Given an input electric driving field $\mathbf{E}(t, \mathbf{r})$ the polarization response of the system is defined through

$$\mathbf{P}(t, \mathbf{r}) = \int_{-\infty}^{+\infty} \chi(t - t')\mathbf{E}(t', \mathbf{r})dt', \tag{4.4}$$

where the response function $\chi(t-t')$ is identically zero for $t < t'$ as required by causality. We should stress that nonlocal integration over the space variable

is not needed here in view of the local (nearest-neighbor) character of the lattice interactions. In a model in which the assumption of locality is relaxed, as for instance in models permitting particle's Lévy flights, the integration over the space variable is expected to produce a physically nontrivial effect. We do not consider such models here. A Fourier transformed $\chi(t)$ defines the frequency-dependent complex susceptibility of the system, $\chi(\omega)$. In a basic theory of polarization response one also introduces the frequency-dependent complex ac conductivity, $\sigma_{\rm ac}(\omega)$, which is related to $\chi(\omega)$ by the Kramers-Kronig integral (Eq. 4.10 below). The dependence of the ac conductivity on frequency, specialized to the random walks on percolation systems, is given by the scaling relation $\sigma_{\rm ac}(\omega) \propto \omega^{\eta}$, where the power exponent η $(0 \leq \eta \leq 1)$ is expressible in terms of the percolation indices β, ν, and μ as $\eta = \mu/(2\nu + \mu - \beta)$ (Gefen et al. 1983). We should stress that the scaling $\sigma_{\rm ac}(\omega) \propto \omega^{\eta}$ incorporates conductivity responses from all clusters at percolation including those finite. In the DPRW model these implications are matched by the mechanism of the hole conduction permitting the polarization current on both infinite and finite clusters. The general linear-response theory expression for the conductivity $\sigma_{\rm ac}(\omega)$ in terms of the mean-square displacement from the origin $\left\langle \mathbf{r}^2(t) \right\rangle$ is (Scher and Lax 1973)

$$\sigma_{\rm ac}(\omega) = \frac{ne^2}{k_B T} D(\omega), \tag{4.5}$$

where

$$\frac{1}{n_d} D(\omega) = \lim_{\epsilon \to 0+} \left[(i\omega)^2 \int_0^{\infty} e^{-i\omega t} e^{-\epsilon t} \left\langle \mathbf{r}^2(t) \right\rangle dt \right] \tag{4.6}$$

with n_d a constant depending on the dimensionality of the lattice and n and e the density and charge of the carriers, respectively. The function $D(\omega)$ has the sense of the frequency-dependent diffusion coefficient (Lax 1958). In the zero-frequency limit, Eq. (4.5) reproduces the well-known Einstein relation between the static diffusion coefficient on the infinite cluster, D_{∞}, and the dc conductivity, $\sigma_{\rm dc} = \lim_{\omega \to 0} \sigma_{\rm ac}(\omega)$. Note that the dc conductivity occurs only through the infinite cluster $(p > p_c)$, as opposed to the ac conductivity response, which occurs through both finite and infinite clusters. In what follows, we require that the frequency ω be large compared to the characteristic evolution frequency in the distribution of the conducting sites. Denoting the latter frequency by ω_*, we have $\omega \gg \omega_*$. Within the present orderings, the scaling of the frequency-dependent diffusion coefficient is given by $D(\omega) \propto \omega^{\eta}$. Consistently with the above definitions, the inverse frequency, $1/\omega_*$, is ordered as the characteristic diffusion time, τ_*, on the infinite cluster, i.e., $\tau_* \simeq \xi^2/D(\omega_*)$. Note that this time will depend on ω_* in accordance with Eq. (4.5). Hence, $\omega_* \simeq 1/\tau_* \simeq D(\omega_*)/\xi^2$, where $\xi \propto |p - p_c|^{-\nu}$ is the diverging pair connectedness length; p is the probability of site occupancy; and p_c is the percolation threshold. We have, at the margins of self-similar behavior, $D(\omega_*) \propto (\omega_*)^{\eta}$, implying that $\omega_* \propto |p - p_c|^{2\nu/(1-\eta)}$. Observe that

$\omega_* \to 0$ for $p \to p_c$. Remembering that there is a microscopic hopping time, which is taken to be unity, we assess the Kubo number in vicinity of the SOC state as

$$Q_* \simeq 1/\omega_* \xi \propto |p - p_c|^{-\nu(1+\eta)/(1-\eta)}. \tag{4.7}$$

One sees that $Q_* \to \infty$ in the limit $p \to p_c$. The Kubo number (Kubo 1963; Brissaud and Frisch 1974) is a suitable dimensionless parameter which quantifies how the evolution processes in the lattice compare with the microscopic hopping motions. The divergency of the Kubo number at criticality shows that there is a time scale separation: fast hopping motions vs. infinitesimal evolution change. In terms of the Kubo number ($Q_* \to \infty$), the diffusion coefficient $D(\omega_*)$ becomes

$$D(\omega_*) \propto \omega_* Q_*^{\gamma}, \tag{4.8}$$

where we have introduced $\gamma = 1 - \eta$. This scaling law appears in models of anomalous diffusion by low-frequency turbulence (Isichenko 1992; Reuss and Misguish 1996; Zimbardo et al. 2000; Pommois et al. 2001; Milovanov 2001, 2009). Special cases of Eq. (4.8) include the well-known Bohm scaling (see, e.g., Dupree 1967), characterized by $\gamma = 1$, as well as the anomalous so-called "percolation" scaling ($\gamma \approx 0.7$), dating back to diffusion-advection models of Isichenko and co-workers (Gruzinov et al. 1990; Isichenko 1991, 1992). Alternatively, the diffusion coefficient on a time varying fractal distribution, Eq. (4.8), can be deduced from the general scaling form (Milovanov 2009)

$$\langle \mathbf{r}^2(t) \rangle = \xi^2(t/\tau_*) f(t/\tau_*), \tag{4.9}$$

where f is a scaling function, which interpolates between the initial-time power-law and flat asymptotic ($t \to +\infty$) behavior: $f(\infty) = \text{const}$. The form in Eq. (4.9) is similar to that considered by Gefen et al. (1983) for anomalous diffusion on percolation clusters (in their model, $\tau_* \propto \xi^{2+\theta}$), and earlier by Straley (1980).

4.4.3 Power-Law Power Spectral Density

By applying the Kramers-Kronig relations $\operatorname{Im} \chi(\omega) \propto \sigma_{\rm ac}(\omega)/\omega$ and

$$\operatorname{Re} \chi(\omega) \propto \text{V.P.} \int \frac{d\omega'}{\omega'(\omega' - \omega)} \sigma_{\rm ac}(\omega') \tag{4.10}$$

it is found that $\chi(\omega) \propto \omega^{-\gamma}$, with $\gamma = 1 - \eta$. A Fourier transformed Eq. (4.4) reads $\mathbf{P}(\omega, \mathbf{r}) = \chi(\omega)\mathbf{E}(\omega, \mathbf{r})$. One can see that the power spectral density, $S(\omega)$, of the system response to a white-noise perturbation, $\mathbf{E}(\omega, \mathbf{r}) = \mathbf{1}$, will be proportional to $|\chi(\omega)|^2$. The end result is summarized as follows:

$$S(\omega) \propto |\chi(\omega)|^2 \propto |\sigma_{\rm ac}(\omega)/\omega|^2 \propto \omega^{-\alpha}, \tag{4.11}$$

where $\alpha = 2(1-\eta) = 2\gamma$. The conclusion is that the power spectral density in the DPRW model is given by an inverse power-law distribution, with the α value depending on scaling properties of the ac conductivity response.

4.4.4 Stretched-Exponential Relaxation and the Distribution of Relaxation Times

Next, we obtain the distribution of relaxation times self-consistently. For this, assume that the system is slightly supercritical, then consider a charge density perturbation, $\delta\rho(t, \mathbf{r})$, caused by the presence of either free charges or holes on the conducting clusters. "Slightly supercritical" means that the dependence of the ac conductivity response on frequency can, with good accuracy, be taken in the power-law form $\sigma_{\rm ac}(\omega) \propto \omega^{1-\gamma}$ discussed above. The implication is that at adding $\delta\rho(t, \mathbf{r})$ to the conducting system at percolation we neglect the departure of the systems geometric properties from pure self-similarity. Without loss in generality, we assume that the perturbation $\delta\rho(t, \mathbf{r})$ is created instantaneously at time $t = 0$. That means that the function $\delta\rho(t, \mathbf{r}) \equiv 0$ for $t < 0$ for all $\mathbf{r}$. The perturbation $\delta\rho(t, \mathbf{r})$ generates an electric field inhomogeneity, $\delta\mathbf{E}(t, \mathbf{r})$ in accordance with Maxwell's equation $\nabla \cdot \delta\mathbf{E}(t, \mathbf{r}) = 4\pi\delta\rho(t, \mathbf{r})$. Consistently with the above discussion, we adopt that for $t > 0$ the decay of $\delta\rho(t, \mathbf{r})$ is due to the spreading of charge-carrying particles (electrons and/or holes) via the random walks on the underlying fractal distribution. The polarization response to $\delta\mathbf{E}(t, \mathbf{r})$ is given by

$$\delta\mathbf{P}(t, \mathbf{r}) = \int_{-\infty}^{+\infty} \chi(t - t')\delta\mathbf{E}(t', \mathbf{r})dt', \tag{4.12}$$

where, as usual, $\chi(t - t') \equiv 0$ for $t < t'$. The density of relaxation currents is defined as the time derivative of $\delta\mathbf{P}(t, \mathbf{r})$, i.e.,

$$\delta\mathbf{j}(t, \mathbf{r}) = \frac{\partial}{\partial t} \int_{-\infty}^{+\infty} \chi(t - t')\delta\mathbf{E}(t', \mathbf{r})dt'. \tag{4.13}$$

The continuity implies that

$$\frac{\partial}{\partial t}\delta\rho(t, \mathbf{r}) + \nabla \cdot \frac{\partial}{\partial t} \int_{-\infty}^{+\infty} \chi(t - t')\delta\mathbf{E}(t', \mathbf{r})dt' = 0. \tag{4.14}$$

Taking $\nabla\cdot$ under the integral sign, then eliminating $\delta\mathbf{E}(t, \mathbf{r})$ by means of Maxwell's equation $\nabla \cdot \delta\mathbf{E}(t, \mathbf{r}) = 4\pi\delta\rho(t, \mathbf{r})$, we find, with the self-consistent charge density,

$$\frac{\partial}{\partial t}\left(\delta\rho(t,\mathbf{r})+4\pi\int_{-\infty}^{+\infty}\chi(t-t')\delta\rho(t',\mathbf{r})dt'\right)=0. \tag{4.15}$$

In writing Eqs. (4.14) and (4.15) we have also assumed that $t>0$. We now integrate in Eq. (4.15) to find

$$\delta\rho(t,\mathbf{r})+4\pi\int_{-\infty}^{+\infty}\chi(t-t')\delta\rho(t',\mathbf{r})dt'=g(\mathbf{r}). \tag{4.16}$$

Here, the function $g(\mathbf{r})$ is an arbitrary function of the position vector $\mathbf{r}$, which appears in the derivation as the constant of integration over time. Under the conditions $\chi(t-t')\equiv 0$ for $t<t'$ and $\delta\rho(t,\mathbf{r})\equiv 0$ for $t<0$ for all $\mathbf{r}$, Eq. (4.16) reduces to

$$\delta\rho(t,\mathbf{r})+4\pi\int_{0}^{t}\chi(t-t')\delta\rho(t',\mathbf{r})dt'=g(\mathbf{r}). \tag{4.17}$$

If we allow $t\to+0$, we find that for $\gamma>0$ the integral term on the left-hand-side goes to zero (as $\propto t^{\gamma}$):

$$\lim_{t\to+0}\int_{-\infty}^{+\infty}\chi(t-t')\delta\rho(t',\mathbf{r})dt'=\lim_{t\to+0}\int_{0}^{t}\chi(t-t')\delta\rho(t',\mathbf{r})dt'=0, \tag{4.18}$$

from which it is clear that $g(\mathbf{r})=\lim_{t\to+0}\delta\rho(t,\mathbf{r})$. We consider this last condition as the initial condition for the relaxation problem. Essentially the same condition holds in the limit $\gamma\to 0$, provided that $\lim_{t\to+0}$ is taken first. A Fourier transformed Eq. (4.17) reads

$$\delta\rho(\omega,\mathbf{k})+4\pi\chi(\omega)\delta\rho(\omega,\mathbf{k})=g(\mathbf{k})/\omega, \tag{4.19}$$

where $\mathbf{k}$ is position vector in reciprocal space, and $g(\mathbf{k})$ is the Fourier image of $g(\mathbf{r})$. Writing the susceptibility as $\chi(\omega)=\tau_{\lambda}^{-\gamma}\omega^{-\gamma}/4\pi$ with τ_{λ} a time constant it is found that

$$\delta\rho(\omega,\mathbf{k})=\frac{1}{\omega+\tau_{\lambda}^{-\gamma}\omega^{1-\gamma}}g(\mathbf{k}). \tag{4.20}$$

The quantity τ_{λ} has the sense of a lifetime of a perturbation with wavelength λ. We expect that $\tau_{\lambda}\propto\lambda^{z}$ at criticality, where z is a scaling exponent. A derivation of this scaling relation will be given shortly. Separating the variables, we write $\delta\rho(\omega,\mathbf{k})=\varphi(\omega)g(\mathbf{k})$, with

$$\varphi(\omega)=1/(\omega+\tau_{\lambda}^{-\gamma}\omega^{1-\gamma}), \tag{4.21}$$

which we consider as the relaxation function in the frequency domain. On inversion to the time domain, Eq. (4.21) generates the Mittag-Leffler function, $E_{\gamma}[-(t/\tau_{\lambda})^{\gamma}]$, which has series expansion (Mittag-Leffler 1905; reviewed in Metzler and Klafter 2000)

$$E_\gamma[-(t/\tau_\lambda)^\gamma] = \sum_{n=0}^{\infty} \frac{[-(t/\tau_\lambda)^\gamma]^n}{\Gamma(1+\gamma n)}. \tag{4.22}$$

Thus, $\varphi(t) = E_\gamma[-(t/\tau_\lambda)^\gamma]$. One sees that the relaxation to SOC of a slightly supercritical state is described by the Mittag-Leffler function $E_\gamma[-(t/\tau_\lambda)^\gamma]$, and not by a simple exponential function as for standard relaxation. We note in passing that the Mittag-Leffler function is the natural generalization of the exponential function. The latter is included as a special case $\gamma = 1$. For times $t \ll \tau_\lambda$, the Mittag-Leffler function, Eq. (4.22), can be approximated by a stretched-exponential the so-called Kohlrausch-Williams-Watts (KWW) relaxation function (Kohlrausch 1854; Williams and Watts 1970)

$$E_\gamma[-(t/\tau_\lambda)^\gamma] \simeq \exp[-(t/\tau_\lambda)^\gamma/\Gamma(1+\gamma)], \tag{4.23}$$

which is often found empirically in various amorphous materials as for instance in many polymers and glass-like materials near the glass transition temperature (for reviews see, e.g., Phillips 1996 and Kaatz et al. 1996; references therein). The KWW relaxation function can conveniently be considered as a weighted average of the ordinary exponential functions, each corresponding to a single relaxation event in the system (Montroll and Bendler 1984):

$$\exp[-(t/\tau_\lambda)^\gamma/\Gamma(1+\gamma)] = \int_0^\infty e^{-t/\Delta t} w_\gamma(\Delta t) d\Delta t. \tag{4.24}$$

The weighting function $w_\gamma(\Delta t)$ is given by Eqs. (51d) and (55) of Montroll and Bendler (1984) where one replaces the exponent α with γ, the time constant T with τ_λ, and the variable μ with $\tau_\lambda/\Delta t$. In our notation:

$$w_\gamma(\Delta t) = (\tau_\lambda/\Delta t^2) L_{\gamma,-1}(\tau_\lambda/\Delta t), \tag{4.25}$$

where $L_{\gamma,-1}$ is the Lévy distribution function with skewness -1 (e.g., Wolfgang and Baschnagel 1999). Assuming a long-wavelength perturbation (i.e., the parameter λ being much longer than the microscopic lattice distance: $\lambda \gg 1$), and setting $\tau_\lambda/\Delta t \gg 1$, we can further approximate the Lévy distribution $L_{\gamma,-1}$ by the Pareto inverse-power distribution. This gives $L_{\gamma,-1}(\tau_\lambda/\Delta t) \propto (\tau_\lambda/\Delta t)^{-(1+\gamma)}$ leading to a pure power-law distribution of relaxation times, consistently with the expectation for SOC:

$$w_\gamma(\Delta t) \propto \Delta t^{-2} \Delta t^{1+\gamma} \propto \Delta t^{-\eta}. \tag{4.26}$$

This power-law distribution was earlier conjectured for SOC on the base of scaling arguments (Tang and Bak 1988). Our conclusion so far is that the relaxations are multi-scale, in accordance with Eq. (4.24), and that their durations are power-law distributed. The distribution is heavy-tailed in the sense that $\int d\Delta t w_\gamma(\Delta t) \propto \tau_\lambda^\gamma \to \infty$ for $\tau_\lambda \to \infty$.

4.4.5 Consistency Check

In a basic theory of dielectric relaxation one writes the frequency-dependent complex dielectric parameter as (Montroll and Bendler 1984; Williams 1989)

$$\epsilon(\omega) - 1 \propto -\int_0^\infty \frac{d\varphi(t)}{dt} e^{i\omega t} dt, \tag{4.27}$$

where $\varphi(t)$ is the relaxation function that describes the decay of polarization after the polarizing electric field has been stepped down or removed instantaneously. In the DPRW model, a step-down type electric field occurs as a consequence of re-injection of the free charges to the infinite cluster. The ensuing relaxation dynamics are mimicked by the chain reactions of hole hopping which act as to properly redistribute the polarization charges across the lattice. Based on the above analysis, we identify the relaxation function in Eq. (4.27) with the Mittag-Leffler function to yield $\varphi(t) \simeq E_\gamma[-(t/\tau_\lambda)^\gamma]$. In vicinity of the critical state, because the upper limit on τ_λ diverges, we can, moreover, replace $E_\gamma[-(t/\tau_\lambda)^\gamma]$ with $\exp[-(t/\tau_\lambda)^\gamma/\Gamma(1+\gamma)]$ for (almost) all $0 < t \leq \infty$. Thus, for $p \to p_c$, $\varphi(t) \simeq \exp[-(t/\tau_\lambda)^\gamma/\Gamma(1+\gamma)]$. Integrating by parts in Eq. (4.27), after a simple algebra one obtains

$$\epsilon(\omega) - 1 \propto 1 - s\mathrm{V}(s) + is\mathrm{Q}(s), \tag{4.28}$$

where $s = \omega\tau_\lambda$ is a dimensionless frequency, and $\mathrm{Q}(s)$ and $\mathrm{V}(s)$ are the Lèvy definite integrals:

$$\mathrm{Q}(s) = \int_0^{+\infty} \exp(-u^\gamma) \cos(us) du, \tag{4.29}$$

$$\mathrm{V}(s) = \int_0^{+\infty} \exp(-u^\gamma) \sin(us) du. \tag{4.30}$$

In the parameter range of multi-scale relaxation response, $\tau_\lambda/\Delta t \gg 1$, $\omega\tau_\lambda \gg 1$, the following series expansions of the Lèvy integrals hold (Montroll and Bendler 1984):

$$\mathrm{Q}(s) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{s^{n\gamma+1}} \frac{\Gamma(n\gamma+1)}{\Gamma(n+1)} \sin\frac{n\gamma\pi}{2}, \tag{4.31}$$

$$\mathrm{V}(s) = \sum_{n=0}^{\infty} (-1)^{n} \frac{1}{s^{n\gamma+1}} \frac{\Gamma(n\gamma+1)}{\Gamma(n+1)} \cos\frac{n\gamma\pi}{2}. \tag{4.32}$$

From Eqs. (4.31) and (4.32) one can see that the expansion of $\mathrm{Q}(s)$ starts from a term which is proportional to $s^{-(1+\gamma)}$, and so does the expansion of $\mathrm{V}(s) - 1/s$. Hence, up to higher order terms, $\epsilon(\omega) - 1 \propto s^{-\gamma}$. Given this, one applies the Kramers-Kronig relations $s\mathrm{Q}(s) \propto \sigma_{\mathrm{ac}}(\omega)/\omega$ and

$$1 - s\mathrm{V}(s) \propto \mathrm{V.P.} \int \frac{d\omega'}{\omega'(\omega' - \omega)} \sigma_{\mathrm{ac}}(\omega') \tag{4.33}$$

to find the scaling of the ac conduction coefficient to be $\sigma_{\mathrm{ac}}(\omega) \propto \omega^{1-\gamma}$. By comparing this with the above expression $\sigma_{\mathrm{ac}}(\omega) \propto \omega^{\eta}$ one reiterates that $\gamma = 1 - \eta$ consistently with the distribution of durations of relaxation events, Eq. (4.26).

4.4.6 Fractional Relaxation and Diffusion Equations

As was shown by Glöckle and Nonnenmacher (1993), the Mittag-Leffler function, Eq. (4.22), is the solution of the fractional relaxation equation

$$\tau_\lambda^\gamma \frac{d\varphi(t)}{dt} = -{}_0D_t^{1-\gamma}\varphi(t), \tag{4.34}$$

where

$${}_0D_t^{1-\gamma}\varphi(t) = \frac{1}{\Gamma(\gamma)} \frac{\partial}{\partial t} \int_0^t \frac{\varphi(t')}{(t-t')^{1-\gamma}} dt' \tag{4.35}$$

is a fractional time the so-called Riemann-Liouville derivative (Podlubny 1999; Metzler and Klafter 2000). Partial cases of this derivative are the unity operator for $\gamma \to 1$ and $\partial/\partial t$ for $\gamma \to 0$. The subscript "zero", added to ${}_0D_t^{1-\gamma}$, signifies that the integration over time on the right-hand-side of Eq. (4.35) starts from $t = 0$. In effect we choose this as the beginning of the system's time evolution. Mathematically, the Riemann-Liouville derivative has the structure of an ordinary time derivative, $\partial/\partial t \equiv \partial_t$, acting on a Laplace convolution of the function $\varphi(t)$ with a power-law, i.e., ${}_0D_t^{1-\gamma}\varphi(t) = [1/\Gamma(\gamma)]\,\partial_t \left[t^{\gamma-1} * \varphi(t)\right] \equiv [1/\Gamma(\gamma)]\, t^{\gamma-1} *' \varphi(t)$, where the symbol $*'$ means that the time differentiation, ∂_t, applies to the entire convolution integral.

It is noticed, following Sokolov et al. (2002), that the Mittag-Leffler function $E_\gamma[-(t/\tau_\lambda)^\gamma]$ describes the relaxation toward equilibrium of particles governed by the fractional diffusion equation, or FDE

$$\frac{\partial}{\partial t} P(t, \mathbf{r}) = {}_0D_t^{1-\gamma} \nabla^2 P(t, \mathbf{r}), \tag{4.36}$$

where $P(t, \mathbf{r})$ is the probability density of finding a particle (random walker) at time t at point $\mathbf{r}$, and the Laplacian operator stands for the local (nearest-neighbor) character of the lattice interactions. Fractional equations of the diffusion type generalize Fick's second law and the Fokker-Planck equation by taking into account memory effects and their occurrence in the SOC problem sounds in fact very natural. The fractional approach is advantageous, as it makes it possible to describe complex systems with anomalous behavior in

much the same way as simpler systems (Metzler and Klafter 2000; Sokolov et al. 2002).

4.4.7 Derivation of the Fractional Diffusion Equation

It is instructive to obtain the fractional diffusion equation (Eq. 4.36) directly from the DPRW relaxation model. For this, let us introduce the electrostatic potential, $\delta\Phi(t,\mathbf{r})$, corresponding to the electric field inhomogeneity, $\delta\mathbf{E}(t,\mathbf{r}) = -\nabla\delta\Phi(t,\mathbf{r})$. Upon substituted into Eq. (4.14),

$$\frac{\partial}{\partial t}\delta\rho(t,\mathbf{r}) = \frac{\partial}{\partial t}\int_{-\infty}^{+\infty}\chi(t-t')\nabla^2\delta\Phi(t',\mathbf{r})dt'. \tag{4.37}$$

In the vicinity of self-organized critical state, we can represent the total charge density, $\rho(t,\mathbf{r})$, as a sum of "unperturbed" or background density, $\rho_c = |e|p_c$, and a perturbation, $\delta\rho(t,\mathbf{r})$, describing the deviation from criticality: $\rho(t,\mathbf{r}) = \rho_c + \delta\rho(t,\mathbf{r})$. To obtain the dependence of $\rho(t,\mathbf{r})$, we use the effective-medium approximation (Bruggeman 1935), a standard technique for calculating average physical properties of many-body systems. The idea is to think of the particles as embedded into an "effective" potential, $\delta\Phi(t,\mathbf{r})$, where they will be Boltzmann-distributed in accordance with

$$\rho(t,\mathbf{r}) = \rho_c \exp\left[e\delta\Phi(t,\mathbf{r})/T\right]. \tag{4.38}$$

Self-consistently, one requires that, on the average, the embedding in the effective medium has the same overall property as the effective medium itself (Bruggeman 1935; Dyre and Schrøder 2000). In writing Eq. (4.38) we took into account that the perturbation, $\delta\rho(t,\mathbf{r})$, is due to negatively charged particles (electrons and/or holes). The normalization condition is defined through $\rho(t,\mathbf{r}) \to \rho_c$ for $\delta\Phi(t,\mathbf{r}) \to 0$. In the above, the parameter T has the sense of a "thermodynamic temperature," associated with the random motion of current-carrying particles on the conducting clusters. For $|e\delta\Phi(t,\mathbf{r})|/T \ll 1$, expanding the exponential function on the right of Eq. (4.38), one finds $\delta\rho(t,\mathbf{r}) \approx (p_c e^2/T)\delta\Phi(t,\mathbf{r})$. Thus, for small perturbations (high temperatures), $\delta\rho(t,\mathbf{r})$ is proportional to $\delta\Phi(t,\mathbf{r})$, as it should. Eliminating $\delta\Phi(t,\mathbf{r})$ in Eq. (4.37), we have

$$\frac{\partial}{\partial t}\delta\rho(t,\mathbf{r}) = \frac{\partial}{\partial t}\int_{-\infty}^{+\infty}\chi(t-t')(T/p_c e^2)\nabla^2\delta\rho(t',\mathbf{r})dt'. \tag{4.39}$$

The memory function, $\chi(t)$, is obtained as Fourier inversion of $\chi(\omega) = \tau_\lambda^{-\gamma}\omega^{-\gamma}/4\pi$, yielding $\chi(t) \propto t^{\gamma-1}$. Under the conditions $\chi(t-t') \equiv 0$ for $t < t'$ and $\delta\rho(t,\mathbf{r}) \equiv 0$ for $t < 0$ for all $\mathbf{r}$, the improper integration in Eq. (4.39) can be performed in the limits from 0 to t. Collecting all dimensional and numer-

ical parameters in one effective "diffusion coefficient," $A_\gamma \propto (T/4\pi p_c e^2)\tau_\lambda^{-\gamma}$, we can write

$$\frac{\partial}{\partial t}\delta\rho(t,\mathbf{r}) = \frac{1}{\Gamma(\gamma)}\frac{\partial}{\partial t}\int_0^t \frac{A_\gamma}{(t-t')^{1-\gamma}}\nabla^2\delta\rho(t',\mathbf{r})dt', \tag{4.40}$$

which is an equivalent form of Eq. (4.36) above. The fractional diffusion equation (Eq. 4.40) can be thought of as deriving from the generalized "Fick's law"

$$\delta\mathbf{j}(t,\mathbf{r}) = -\frac{1}{\Gamma(\gamma)}\frac{\partial}{\partial t}\int_0^t \frac{A_\gamma}{(t-t')^{1-\gamma}}\nabla\delta\rho(t',\mathbf{r})dt', \tag{4.41}$$

or $\delta\mathbf{j}(t,\mathbf{r}) = -{}_0D_t^{1-\gamma}A_\gamma\nabla\delta\rho(t,\mathbf{r})$, which is readily deduced from Eq. (4.13) in the effective-medium approximation. One sees that the current $\delta\mathbf{j}$ involves in some way the past values of the concentration gradient, so that the dependence is not instantaneous in general, by contrast with the traditional Fickian case. Equation (4.41) can equivalently be obtained from the general scaling law for anomalous diffusion on percolation systems (Milovanov 2009). A derivation using CTRW's can be found in Metzler et al. (1998, 1999). We should stress that the non-Markovian nature of Eqs. (4.40) and (4.41), together with the fractional relaxation equation (Eq. 4.34) accounts for the long-time memory effects in SOC phenomena, where one believes the time evolution exhibits long tails and infinite correlation scale (Sornette 1992a).

The occurrence of the fractional diffusion equation (Eq. 4.40) might be interpreted, with the aid of the proposed SOC model, in favor of considering SOC as one important case for *fractional kinetics* (Shlesinger et al. 1993). Indeed the concept of fractional kinetics enters different areas of research, such as turbulent transport in plasmas and fluids, particle dynamics in potential fields, quantum optics, and many others. This subject is summarized in comprehensive reviews (Metzler and Klafter 2000, 2004; Zaslavsky 2002). In many ways equations built on fractional derivatives offer an elegant and powerful tool to describe anomalous transport in complex systems (Zaslavsky 2002). There is an insightful connection with a generalized master equation formalism along with a mathematically convenient way for calculating transport moments as well as solving initial and boundary value problems (Metzler and Klafter 2000, 2004). The fundamental solution or Green's function of the fractional Eq. (4.36) is evidenced in Table 1 of Metzler and Klafter (2004).

4.4.8 Dispersion-Relation Exponent

In sandpile SOC models, one is interested in how the lifetime of an activation cluster scales with its size (Zhang 1989). In the DPRW model, by activation cluster one means a connected cluster of activated sites. An occupied site is said "activated" if it has become a hole or if it contains a free charge. Clearly,

activation clusters can only exist above the percolation threshold. Note that activation clusters are subsets of the underlying conducting cluster of polarization charges. The notion of activation cluster is but a visualization of the charge density inhomogeneity $\delta\rho(t,\mathbf{r})$ in terms of a connected distribution of activated sites. Activation clusters decay because the constituent charged particles (holes and/or free charges) diffuse away via the random walks.

Consider an isotropic activation cluster composed of free particles. (The nature of the particles does not matter here − the hole case is just similar.) It is assumed for convenience, without loss of generality, that each site of an activation cluster contains only one particle. Thus, the number density of the free particles inside the activation area is equal to one. It steps down to zero just outside. If the microscopic lattice distance is a ($a = 1$), then there is a unit density gradient across the boundary of the activation cluster looking inside. Because of this gradient the activation cluster will be loosing particles on the average. A particle that has crossed the boundary against the direction of the gradient is considered to be lost from the cluster. As the particles dissipate, the location of the boundary shifts inward with speed u. The local flux density of those particles leaving per second the activation area is just the gradient times the local diffusion coefficient. The latter depends on frequency of the relaxation process as $D(\omega) \propto \omega^{\eta}$ in accordance with Eq. (4.5). If l is the current size of the cluster, then the corresponding relaxation frequency is $\omega \simeq u/l$. Using this, the frequency dependence of the diffusion coefficient can be translated into the corresponding l-dependence, the result being $D(l) \propto l^{-\eta}$. Balancing the rate of decay of the cluster with the outward flux of the particles we write $dl/dt \propto -l^{-\eta}$. Integrating this simple equation over time from $t = 0$ to $t = \tau_\lambda$ and over l from $l = \lambda$ to $l = 0$ one finds the dispersion relation $\tau_\lambda \propto \lambda^z$, typical for the SOC phenomena, with the exponent $z = 1+\eta = 2-\gamma$.

4.4.9 The Hurst Exponent

The persistency of relaxation is measured by the Hurst exponent, H, which is related to our $z = 1 + \eta$ via $H = 1/z$. The Hurst exponent finds its significance in the statistics of self-affine graphs of the signals where it defines the correlations between past and future increments (Hurst 1951; Mandelbrot and Van Ness 1968; Feder 1988). Depending on the value of H, one distinguishes between fractional Brownian motions ($0 < H < 1$) and fractional Brownian noises ($H < 0$). (See, also, section 2.3.3, this book.) The difference is that a noise-like function has a stationary quality in the sense that its variance ($\propto t^{2H}$) does not asymptotically grow with time. We note in passing that $H = 0$ is a very special value which separates motion-like and noise-like response processes. Focusing on the motion-like signals, $H > 1/2$ represents persistence (super-diffusion), and $H < 1/2$ anti-persistence (sub-diffusion). Noting here that the exponent η takes the values between 0 and

1, one sees that the value of $H = 1/(1+\eta)$ lies between 1/2 and 1. In this parameter range, the signal is motion-like persistent. That means, precisely, that the existing tendency is maintained through the dynamics (Feder 1988). Thus, the correlations are present at all time scales, consistently with the expectation for SOC. It is in this sense that a persistent signal is said to be long-time correlated (Feder 1988). In the mean-field (i.e., diffusion) limit of $z \to 2$, the correlation vanish, so that the usual "Brownian" value, $H = 1/2$, is reinstalled.

4.4.10 Activation-Cluster Size Distribution and the τ-Exponent

Following Zhang (1989), we introduce the distribution law of the size ς of activation clusters, $w_\tau(\varsigma) \propto \varsigma^{-\tau+1}$, which also defines the distribution of the particle flows caused by a single chain reaction (Bak et al. 1987, 1988; Tang and Bak 1988).[8] The exponent of the power-law, τ, is obtained from Eq. (5) of Tang and Bak (1988), where one replaces the "noise" exponent ϕ with α in accordance with Eq. (4.11) above, and the fractal dimension D with the hyperscaling $d_f = d - \beta/\nu$, leading to $\tau = 3 - \alpha z/d_f$. The present result differs from the expression obtained in Zhang (1989) in that it takes into account the fractal geometry of activation clusters in the vicinity of the criticality. Note that the dependence on the ambient dimension, d, enters the τ value through the hyperscaling, i.e., $\tau = 3 - \alpha z/(d - \beta/\nu)$, which also involves the percolation indices β and ν. When one notes that the exponent of the stretched-exponential relaxation is given by $\gamma = 1 - \eta = 2 - z$ (see Eq. 4.24), one may apply the formula $\tau = 3 - \alpha z/d_f$ in any $d \geq 1$, consistently with the implication of the "hyperuniversal" fractal dimension, $d_f = d - \beta/\nu$ (e.g., Nakayama et al. 1994; Havlin and ben-Avraham 2002), and at odds with the mention in Tang and Bak (1988, p. 2349) that it gives rise to "logarithmic singularities" in $d = 2$.

4.4.11 Occurrence Frequency Energy Distribution and the ß-Exponent

It is convenient to think of the activation clusters as containing a certain amount of "energy" which is released when the comprising free particles dissipate to the boundaries. Using here that the electric charge of the free particles is a conserved quantity, we may associate the energy content of the

[8] Here we use $-\tau + 1$ as in Tang and Bak (1988), instead of $-\tau$ as in Zhang (1989), to be the exponent of $w_\tau(\varsigma)$.

activation clusters with their electric charge content. When one notes that, by assumption, each site of an activation cluster contains only one particle, and that, moreover, the clusters are self-similar at percolation, one finds that the energy confined by a cluster of size Δl is proportional to its fractal volume, $\Delta V_f \propto \Delta l^{d_f}$, where $d_f = d - \beta/\nu$ is the "hyperuniversal" fractal dimension. Thus, we have, for the energy content, $\Delta\epsilon \propto \Delta V_f \propto \Delta l^{d_f}$. More so, the energy confinement time, Δt, is obtained as the activation cluster lifetime. The latter is defined from the dispersion relation to be $\Delta t \propto \Delta l^z$, where z is the dispersion relation exponent. The implication is that each relaxation event with the duration Δt dissipates electric charges at a certain spatial scale, Δl, so that the characteristics of the relaxation events in space and time are inherently coupled in the vicinity of SOC, as they should. Eliminating Δl with the aid of the dispersion relation, we arrive at the scaling $\Delta\epsilon \propto \Delta t^{d_f/z}$, from which the mean-field behavior, $\Delta t \propto \Delta\epsilon^{1/2}$, can be deduced. Differentiating in Eq. (4.26) with respect to Δt, we obtain the occurrence frequency distribution of durations of relaxation events $N_\gamma(\Delta t) = dw_\gamma(\Delta t)/d\Delta t \propto \Delta t^{-\eta-1}$. If we substitute this distribution with the scaling of the confined energy, we find

$$N_\gamma(\Delta\epsilon) = dw_\gamma(\Delta\epsilon)/d\Delta\epsilon = dw_\gamma(\Delta t)/d\Delta t \cdot d\Delta t/d\Delta\epsilon \propto \Delta\epsilon^{-ß}, \qquad (4.42)$$

where the exponent of the power-law is given by $ß = 1 + \eta z/d_f$. The distribution in Eq. (4.42) has the sense of occurrence frequency energy distribution of relaxation events in a dynamical system at SOC. This distribution finds its significance in the study of solar flares and gamma bursts (Hudson 1991; Charbonneau et al. 2001; Aschwanden 2011, 2012; references therein). More applications can be proposed for the statistical physics of faulting and earthquakes (e.g., Jensen 1998, Turcotte 1999). Indeed the distribution in Eq. (4.42) can be interpreted as a power-law distribution of seismic energy released through earthquakes, equivalent to the Gutenberg-Richter frequency-magnitude law (Gutenberg and Richter 1954; section 5.1, this book), with significant implications for practical problems in probabilistic seismic hazard evaluation (Main 1996). We estimate the ß values shortly.

4.4.12 Values of the Critical Exponents

Using known estimates (Stauffer, 1979; Isichenko 1992; Nakayama et al. 1994) of the percolation indices β, ν, and μ we could evaluate the critical exponents of the DPRW model in all ambient dimensions $d \geq 1$. The results of this evaluation, summarized in Table 4.2, are in good agreement with the reported numerical values from the traditional sandpiles (for $d = 2$, $z \approx 1.29$, $\tau \approx 2.0$; for $d = 3$, $z \approx 1.7$, $\tau \approx 2.33$) (Tang and Bak 1988) and earlier theoretical predictions (for $d = 2$, $z = 4/3$, $\tau = 2$; for $d = 3$, $z = 5/3$, $\tau = 7/3$) (Zhang 1989). We consider this conformity as a manifestation of the universality class

of the model.[9] For $d = \infty$, the model reproduces the exponents of mean-field SOC (see, e.g., Vespignani and Zapperi 1998).

In many ways, the DPRW approach to SOC offers a simple yet relevant lattice model for dielectric relaxation phenomena in systems with spatial disorder. One by-product of this approach is a consistent theoretical derivation of the KWW stretched-exponential relaxation function, Eq. (4.23). In this respect, we observe that the model gives values of the exponent γ (for $d = 2$, $\gamma \approx 0.66$; for $d = 3$, $\gamma \approx 0.4$) in good agreement with the typical experimental results (the γ value between 0.3 and 0.8) (Montroll and Bendler 1984; Phillips 1996; Capaccioli et al. 1998; Jacobs et al. 2006). This observation supports the hypothesis (Milovanov et al. 2007, 2008) that dielectrics exhibiting stretched exponential relaxations are in a state of self-organized criticality.

More so, the DPRW model gives a Hurst exponent (for $d = 2$, $H \approx 0.75$; for $d = 3$, $H \approx 0.6$) consistently with the reported narrow range of variation of H as observed in different magnetic confinement systems (Hurst exponent varying between $H \approx 0.62$ and 0.75) (Carreras et al. 1998, 1999, 2001; Pedrosa et al. 1999). In this connection, it is worth adding here that SOC behavior of the bulk plasma transport is expected to be a characteristic of higher-power plasma discharges in the so-called low confinement regime (Carreras et al. 1998).

With respect to the occurrence frequency energy distribution, Eq. (4.42), the model predicts that ß = 1 in one dimension (for $d = 1$, $\eta = 0$) and ß = 3/2 in the mean-field limit (for $d \geq 6$, $z = 2$, $\eta = 1$, and $d_f = 4$). These results are exact. Also, one finds, approximately, ß ≈ 1.24 for $d = 2$ and ß ≈ 1.4 for $d = 3$ (Hausdorff fractal dimensions $d_f = 91/48$ and $d_f \approx 2.5$, accordingly) (see Table 4.2). These theoretical predictions are in close agreement with the predictions of Aschwanden (2012) who found values of ß = 1.0 for $d = 1$; ß = 1.28 for $d = 2$; and ß = 1.5 for $d = 3$ from a statistical fractal-diffusive avalanche model of a slowly-driven self-organized criticality system (section 2.2.2, this book. In the notation of Aschwanden 2012, ß = α_E). The latter model assumes that the avalanche size grows as a diffusive random walk, implying that $z = 2$ in all dimensions $d = 1, 2, 3$. In our model, diffusive random walks are recovered in the mean-field limit only and in relatively high ambient dimensions that are not smaller than 6, while the z exponent is taken to be non-diffusive in general. Even so, this does not seem to affect the ß value significantly, so that our mean-field result, ß = 3/2, almost precisely coincides with the result of Aschwanden (2012) in three dimensions. Earlier, Litvinenko (1998) has suggested that the distribution of flare energies is characterized by a power-law with the slope ß = 3/2 independently of the ambient dimensionality $d > 1$. He modeled an avalanching process on a tree without loops, thus giving rise to this value. In this context, we should stress that the effect of loops can be abandoned only in the high dimensions

[9] We should stress that we observe this numerical conformity despite that the underlying analytical expressions are fairly different. Compare with Tang and Bak (1988) and Zhang (1989).

$d \geq 6$, permitting a mean-field description (Nakayama et al. 1994; Havlin and ben-Avraham 2002). All in all, the exponent ß = 3/2 agrees well with the reported slopes of the occurrence frequency energy distribution for solar flares (around −1.5 to −1.8) (Hudson 1991; Crosby et al. 1993; Wheatland and Uchida 1999), demonstrating that the observed power-law distribution of flare energy release is well reproduced under the assumption that the solar corona operates as a self-organized criticality system (Vlahos et al. 1995; Charbonneau et al. 2001; Norman et al. 2001; Aschwanden 2011, 2012).

Finally, it is worth noting here that exponents similar to the above (around −1.5 to −1.8) are also found in earthquake phenomenology, with some geographical dependence in the ß value (Main 1996; Jensen 1998; Turcotte 1999; references therein). In this regard, a differential Gutenberg-Richter distribution of seismic energies with the exponent ß = 3/2 was proposed by Sornette (1992b) from a mean-field version of the Burridge-Knopoff block-spring model (e.g., Main 1996) of earthquakes. This similarity with the statistical physics of faulting and earthquakes indicates that the behavior is relatively insensitive to the details of the system that is analyzed. By contrast, there is a stronger ingredient due to the general properties of heavy-tailed occurrence frequency distributions in the vicinity of the criticality as dictated by the multi-scale relaxation dynamics of the Mittag-Leffler type, and by associate fractional relaxation equation (Eq. 4.34). It is this ingredient that we believe to characterize the operation of systems with many interacting degrees of freedom (i.e., "complex" systems), thus rewarding the notion of SOC.

Table 4.2: Critical exponents of the DPRW model. Exponents appearing in statistical distributions as for instance inverse power-law power spectral density distribution are summarized in the lower part of the table. The mean-field results, holding for $d \geq 6$, are collected as $d = \infty$. Input parameters are the percolation indices β, ν, and μ (Stauffer, 1979; Isichenko 1992; Nakayama et al. 1994).

Exponent	Expression	Description	$d = 1$	$d = 2$	$d = 3$	$d = \infty$
η	$\mu/(2\nu + \mu - \beta)$	ac conductivity	0	0.34	0.6	1
z	$1 + \eta$	Dispersion relation	1	1.34	1.6	2
γ	$1 - \eta$	Mittag-Leffler relaxation	1	0.66	0.4	0
H	$1/z$	Hurst exponent	1	0.75	0.6	1/2
Exponent	Expression	Distribution	$d = 1$	$d = 2$	$d = 3$	$d = \infty$
η	$\mu/(2\nu + \mu - \beta)$	Relaxation-time	0	0.34	0.6	1
α	$2 - 2\eta$	Power spectral density	2	1.3	0.8	0
τ	$3 - \alpha z/d_f$	Activation-cluster size	1	2.1	2.5	3
ß	$1 + \eta z/d_f$	Occurrence frequency energy	1	1.24	1.4	3/2

4.5 The Random Walk's Guide to SOC

Let us take stock at this point and address a few important issues concerning the DPRW model of SOC and its random-walk ingredient.

4.5.1 General

Apart from details of the mathematical formalism, the DPRW model is actually quite simple. The main points are as follows. A lattice site can either be empty or occupied. An occupied site is interpreted as polarization charge. The equilibrium concentration of the polarization charges depends on the potential difference between the plates. When the potential difference changes the lattice occupancy parameter adjusts. A dynamical mechanism for this uses holes. The holes are just missing polarization charges. They are important key elements to the model as they provide a mechanism for the polarization current in the system. Beside holes, the free charges are introduced. The free charges, too, carry electric currents whose very specific role in the model is just to control the potential difference between the plates. The changing amount of the free charges in the system has an effect on the lattice occupancy parameter. Nonlinearly, it affects the conductivity of the lattice. This nonlinear twist provides a dynamical feedback by which the system is stabilized at the state of critical percolation. In many ways the proposed model is but a simple lattice model for dielectric relaxation in a self-adjusting disordered medium. It is perhaps the simplest model which accounts for the whole set of relaxation processes including the hole conduction.

It is worth assessing the advantages and disadvantages of the DPRW approach to SOC. In terms of advantages, the electric nature of the model greatly facilitates the analytical theory: Not only does it permit to quantify the microscopic lattice rules in terms of the frequency-dependent complex ac conductivity, the use of the Kramers-Kronig relation in Eq. (4.10) makes it possible to directly obtain the susceptibility function by integrating the conductivity response. As a result, the exponents z, γ, α, and H are expressible in terms of only one parameter, the exponent of ac conduction η. The latter is obtained as a simple function of the percolation indices β, ν, and μ.

With respect to disadvantages, the model is seemingly different from the traditional approaches to SOC based on cellular automation (CA) and its integration in the existing family of SOC models might be a matter of debate. Even so, the idea of the random walks on a self-organized percolation system as a simplified yet relevant model for SOC constitutes a significant appeal: First, it relies on the established mathematical formalism of the random walks (Bouchaud and Georges 1990; ben-Avraham and Havlin 2000; Havlin and ben-Avraham 2002; Metzler and Klafter 2000, 2004) whose advance on the SOC problem is theoretically very beneficial. Second, it offers a clear

connection to studies outside the conventional SOC paradigm as for instance to transport of mass and charge in disordered media (Lax 1958; Scher and Lax 1973; Druger et al. 1985; Dyre and Schrøder 2000). Instead, the traditional CA type models are complicated by a poor analytical description of the microscopic transport mechanisms and their basic physics appreciation is at times uneasy.

4.5.2 The Role of Random Walks

It is theoretically important to note that the dielectric context of the considered model, apart from offering a convenient platform for the analytical theory, is actually not essential for the SOC phenomena. Indeed the DPRW model could be defined in terms of diffusion processes for neutral particles of different kinds. A formal reason for this is the equivalence (Lax 1958; Scher and Lax 1973) of the frequency-dependent electrical conductivity problem and the frequency-dependent diffusion problem, specific to hopping conduction. The crucial element to the model is in fact the assumption of the random walks, not the nature of the particles.

The possible generalizations of the DPRW model correspond to biased random walks of the free charges in the direction of the potential drop and/or inclusion of a second critical threshold $p_{cc} \geq p_c$ above which the random walk dynamics might change to a biased motion. We consider those generalizations obvious as they mainly intend to modify the value of the exponent η in a certain parameter range, while the basic physical picture of SOC will remain essentially the same.

4.5.3 Universality Class

The final point to be addressed here concerns the issue of universality class. We take notice of the fact that the DPRW SOC model uses the *charitable* redistribution rule (Maslov and Zhang 1996) to propagate the activities, likewise to the traditional BTW sandpile (Bak et al. 1987, 1988) or similar (Tang and Bak 1988; Zhang 1989). That means that an active site always loses its content to the neighbors. The charitable rule is to be distinguished from the *neutral* rule, when each of $2d + 1$ sites involved in redistribution gets an unbiased random share of the transported quantity. Models using the neutral rule often fall in the universality class of directed percolation (DP) and are characterized by appreciably larger values of the dynamic exponent z (for $d = 2$, $z \approx 1.73 \pm 0.05$) (Maslov and Zhang 1996). Based on this evidence, we suggest that the DPRW model belongs to the same universality class as

the BTW sandpile, and not to the DP universality class, consistently with the values of the critical exponents collected in Table 4.2.

4.6 Self-Organized Turbulence: The "Sakura" Model

The DPRW model of SOC can be extended so that it includes the phenomena of self-organized "turbulence" in Earth's geomagnetic tail, discussed in Milovanov et al. (1996, 2001a,b) and Zelenyi et al. (1998). The main idea here is that, when the processes of magnetic reconnection stretch the magnetotail beyond a certain limit, the cross-tail electric current system of Earth's dipole-like magnetic field is destabilized, since it crucially requires the presence of a regular component of the geomagnetic field normal to the current sheet plane. Then the magnetotail spontaneously evolves into a far from equilibrium dynamical "turbulent" state, where it responds to changes in the tail current intensity and the varying dawn-dusk potential difference in terms of turbulent perturbation electric currents and magnetic field fluctuations. It was argued that the transport of electric charge across the magnetotail was due to heavier plasma species, the ions, whose regular orbits (transient or Speiser, weakly trapped, trapped) were essentially destroyed in the absence of stabilizing normal magnetic field and that the steady state which is self-organized corresponded to a strongly shaped electric current system and to a highly inhomogeneous, multi-scale magnetic fluctuation pattern (Fig. 4.5). This model for the coupled turbulent perturbation electric currents and magnetic field fluctuations has come to be known as "Sakura" model, after its presentation at the Chapman Conference held in Kanazawa, Japan, November 5-9, 1996 (see Zelenyi et al. 1998). Applications of the Sakura model pertain to the phenomena of tail current disruption in substorm regions of the near-Earth tail (Milovanov et al. 2001a,b) as well as to the explanation of permanent presence of magnetic field fluctuations in the distant magnetotail (Milovanov et al. 1996; Zelenyi et al. 1998), as suggested by the GEOTAIL measurements (Hoshino et al. 1994; Nishida et al. 1994). The important role of the ion component of the plasma in driving cross-tail current instability was addressed by Lui et al. (1995); Ohtani et al. (1995); and Sharma et al. (2006), where one also finds an analysis of the satellite observational data.

Denoting the perturbation tail current density and magnetic fluctuation field by respectively $\delta\mathbf{j}(t,\mathbf{r})$ and $\delta\mathbf{B}(t,\mathbf{r})$, we write

$$\nabla\times\delta\mathbf{B}(t,\mathbf{r})=\frac{4\pi}{c}\delta\mathbf{j}(t,\mathbf{r}), \tag{4.43}$$

where, under the assumptions of locality and linear conductivity response,

$$\delta\mathbf{j}(t,\mathbf{r})=\int_{-\infty}^{+\infty}\sigma(t-t')\delta\mathbf{E}(t',\mathbf{r})dt'. \tag{4.44}$$

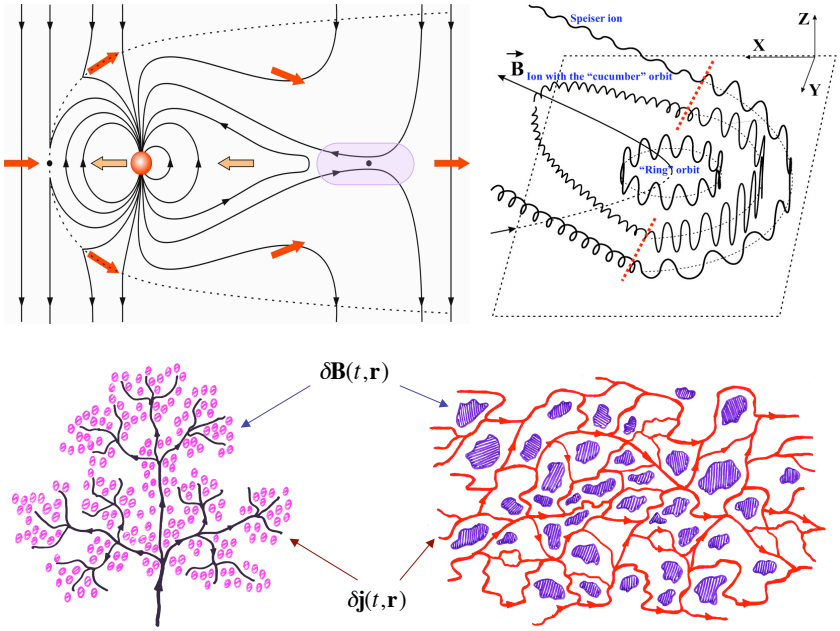


Fig. 4.5: "Sakura" model. Upper left: Schematic illustration for the Earth's dynamic magnetosphere interacting with its solar wind drive. Arrows indicate the magnetic field and plasma convection, where, red arrows represent the convection from the solar wind into the magnetotail, and brown arrows toward Earth and in the direction of the dayside magnetopause. Considerable local stretching and thinning of the magnetotail occurs in the vicinity of the neutral X-line during the substorm growth phase. Similar conditions for stretching and thinning of the Earth's dipole-like magnetic field − associated with the vanishing of the regular component of the geomagnetic field normal to the current sheet plane − are believed to verify in the distant Earth's magnetotail at the geocentric distances larger than $50 - 100$ Earth's radii (not shown here) (Milovanov et al. 1996). Upper right: The various orbit types for thermal ions in the magnetotail current sheet with finite regular component of the normal field. Adapted from Sharma et al. (2006). When the processes of magnetic reconnection stretch the magnetotail beyond a certain limit, the particle regular orbital motion is essentially destabilized and the cross-tail electric current system spontaneously evolves into a far from equilibrium dynamical "turbulent" state where it responds to changes in the tail current intensity and the varying dawn-dusk potential difference in terms of turbulent perturbation electric currents and magnetic field fluctuations. Bottom left: Original model promotion as a blossoming sakura tree: Its "leaves" represent the magnetic field fluctuations; its "branches," the perturbation electric currents. Bottom right: Plan view of the turbulent current sheet. Cross-tail electric currents are organized in highly branched, very inhomogeneous conducting patterns with the topology of a fractal network at percolation (red color). The magnetic field fluctuations are shown as chunks of different sizes (violet color). They scatter the momentum of current-carrying particles, the ions, and are electromagnetically related with the perturbation electric current intensity by means of Maxwell's equation, $\nabla \times \delta\mathbf{B}(t,\mathbf{r}) = (4\pi/c)\delta\mathbf{j}(t,\mathbf{r})$. Note that the magnetic field fluctuations, $\delta\mathbf{B}(t,\mathbf{r})$, can be thought as analog polarization charges in the DPRW model of SOC.

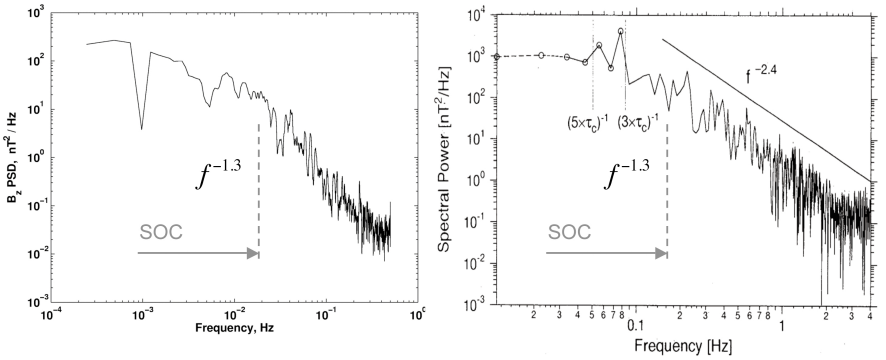


Fig. 4.6: The typical spectra of magnetic field fluctuations in the near-Earth stretched and thinned magnetotail prior to tail current disruption. Left: INTERBALL-1 observations. Adapted from Zelenyi et al. (1998). Right: A spectrum observed by the Charge Composition Explorer of the Active Magnetospheric Particle Tracer Explorers (AMPTE) satellite, with an emphasis on the August 28, 1986, current disruption event. Adapted from Ohtani et al. (1995). A distinctive feature of these spectra is the knee around a characteristic frequency posed by the unstable tearing modes. The spectrum is flatter below the knee frequency and is noticeably steeper just above it. The processes of self-organization to a critical state correspond to the flatter counterpart of the spectrum and to frequencies lesser than the knee frequency: in practice, smaller than $\sim 5 \cdot 10^{-2}$ Hz, although the exact bound, in real data, may not be that certain. Associated spectrum is a power-law with the typical slope ~ -1.3 (in log-log plot). In the limit of very low frequencies, the observed signals cross over to a white noise, as they should (this is due to finite system size effects and/or the natural limitations of the observational time series), so that the spectra are flat. The spectral properties of the fluctuations just above the knee (slope ~ -2.4) involve the processes of convection of magnetic turbulence structures with decelerated solar wind velocity along the magnetotail. These processes were analyzed by Milovanov et al. (2001a,b) and are not considered here. All in all, the signatures of self-organization to a critical state are to be expected in the intermediate frequency range, comprised between those frequencies where the spectra are white noise-like and the knee frequency. An approximate position of the knee is marked by vertical dashed line.

Here, $\delta\mathbf{E}(t,\mathbf{r})$ is the perturbation electric field in the current sheet plane. The memory function, $\sigma(t)$, is obtained as Fourier inversion of the frequency-dependent complex ac conductivity, $\sigma_{\rm ac}(\omega)$, where the ac dependence is due to the turbulent nature of the conducting domain. It is understood that the time varying magnetic perturbation generates a time and spatially varying electric field because of Faraday's law. An important feature which arises

in this induction process is gradual heating and energization of the plasma, leading to the occurrence of slowly decaying, high-energy non-thermal wings in the particle energy distribution function (Milovanov and Zelenyi 2001, 2002; Zelenyi and Milovanov 2004). Often in the magnetospheric plasma research those distributions with wings are modeled by non-thermal the so-called "kappa" distributions (Christon et al. 1989) which interpolate between the initial low-energy exponential forms and the asymptotic inverse power-law behavior. The significance of the "kappa" distributions lies in the fact (Milovanov and Zelenyi 2000) that they appear as canonical distributions in the non-extensive thermodynamics due to Tsallis (1988; section 2.3.5, this book). In the present analysis we shall assume, however, that all fluctuation frequencies are so slow that the effect of the inductive field can be neglected and we omit, consequently, the inductive term in Eq. (4.43) above. To this end, performing a Fourier transform of Eq. (4.43) in space and time, we have

$$\mathbf{k} \times \delta\mathbf{B}(\omega, \mathbf{k}) = \frac{4\pi}{c}\delta\mathbf{j}(\omega, \mathbf{k}), \tag{4.45}$$

where $\mathbf{k}$ is the wave vector of the perturbation. Simultaneously, from Eq. (4.44) we find $\delta\mathbf{j}(\omega, \mathbf{k}) = \sigma_{\rm ac}(\omega)\delta\mathbf{E}(\omega, \mathbf{k})$. In vicinity of the current instability threshold, considering the low-frequency limit of Eq. (4.43), we may think of the fluctuations as of collection of plane waves, characterized by a linear dispersion relation $\omega = \mathbf{k} \cdot \mathbf{u}$, where the phase velocity, $\mathbf{u}$, does not depend on $\mathbf{k}$. Given that the input perturbing electric field is uncorrelated white noise: $\delta\mathbf{E}(\omega, \mathbf{k}) = \mathbf{1}$, with the aid of Eq. (4.45) one sees that the power spectral density of the magnetic fluctuation field, $|\delta\mathbf{B}(\omega, \mathbf{k})|^2$, will be proportional to $S(\omega) \propto |\sigma_{\rm ac}(\omega)/\omega|^2$. At this point, if one assumes, following Milovanov et al. (2001a,b), that the dynamical "turbulent" state is characterized by fractal geometry of the threshold percolation, one may exploit the scaling relation $\sigma_{\rm ac}(\omega) \propto \omega^{\eta}$ to obtain $S(\omega) \propto \omega^{-\alpha}$, with $\alpha = 2(1 - \eta)$, consistently with the DPRW result. Utilizing the percolation estimate above (for $d = 2$, $\eta \approx 0.34$), one finds that $\alpha \approx 1.3$. This theoretical prediction compares well against the reported α values in the lower-frequency part of the magnetic fluctuation spectrum (below a turnover or knee frequency posed by the unstable tearing modes: see Fig. 4.6) (Hoshino et al. 1994; Bauer et al. 1995; Ohtani et al. 1995). Thus, the behavior is self-similar in the self-organization domain. Note that the power spectral density of the magnetic fluctuation field, $|\delta\mathbf{B}(\omega, \mathbf{k})|^2$, corresponds with the power spectral density of the dynamic polarization response, Eq. (4.11). We should stress that the Sakura model leads to a smaller spectral index ($\alpha \approx 1.3$) than Kolmogorov's theoretical value for fluid turbulence ($\alpha = 5/3$) as well as Kraichnan's theoretical value for magnetohydrodynamic plasma turbulence ($\alpha = 3/2$). This last observation addresses the significance of "self-organized" magnetic fluctuation turbulence as opposed to ordinary (fluid-like) turbulence.

Our conclusion so far is that the power spectral density of the magnetic fluctuation field is given by an inverse power-law and that the behavior cor-

responds with the prediction of the DPRW SOC model. This result suggests that micro-processes of self-organization of electric currents and magnetic field fluctuations in the Earth's stretched and thinned magnetotail are governed by SOC (Milovanov et al. 2001b; Zelenyi and Milovanov 2004). Note that the above assumption of fractality and self-similar behavior is validated through the direct analysis of time series of the satellite observed magnetic field fluctuations, obtained *in situ* in the substorm regions of the near-Earth tail and associated with the phenomena of tail current instability during the substorm growth phase (Ohtani et al. 1995).

4.7 Beyond Linear Theories: DANSE Formalism

Theoretical approaches discussed so far were based on a linear-response theory and on fractional generalizations of the diffusion and relaxation equations (Eqs. 4.34 and 4.40), which are linear by construction. Nonlinearities were contained in fractal geometry of the dynamical system at percolation and implicitly in the various "avalanche" exponents and the fractional indices of time differentiation. In the present analysis, this paradigm of "linear dynamics in a nonlinear medium" will be relaxed. Rather, a more general theoretical picture will be drawn, in which the dynamical and structural nonlinearities are twisted, and for which one might propose the formula "nonlinear dynamics in a nonlinear medium." Very specifically, we intend to demonstrate that the phenomena of SOC − at least those belonging to the universality class of the DPRW model − may be cast in the mathematical formalism of discrete Anderson nonlinear Schrödinger equation (DANSE), which invokes a random potential for lattice interactions, and in which the strength of nonlinearity, being an inherent part of the model description, is determined dynamically as the system self-adjusts and evolves to criticality in response to external forcing. Most previous theoretical studies of SOC have neglected the possibility of describing the lattice interactions in terms of a random potential field, and have focused, consequently, on the microscopic redistribution rules for the dynamics. We believe that this approach is unnecessarily restrictive and has left out the important physics results.

4.7.1 The Roadmap

As is already mentioned above, in the DPRW model the critical state is made self-organized via the mechanisms of hole-hopping by which the system responds to the fluctuating potential difference on the capacitor. We should stress that the holes redistribute the polarization charges in a way as to preserve the properties of the random percolation. They change the shape and

the folding of the percolation clusters in the ambient configuration space, but not the random character in the distribution of the conducting sites. A confirmation of these ideas can be obtained if one considers holes as "excitations" of self-organized critical state mediating the lattice activities. With this interpretation in mind, the transport problem for holes can be formulated as a transport problem for the hole wave function in a random potential field.[10] The latter problem, in its turn, can be studied as part of the general problem of transport of waves in disordered media, with that very specific element that the medium is nonlinear and its properties are coupled with the wave process itself − which, too, can be nonlinear.

We consider the problem of dynamical localization of the hole wave function in the framework of nonlinear Schrödinger equation (NLSE) with a random potential term on a lattice. The NLSE was derived for a variety of physical systems under some approximations (Ablowitz and Segur, 1981). Recently, it has been rigorously established that, for a large variety of physical conditions and details of interaction, the NLSE (also known as the Gross-Pitaevsky equation) is exact in the thermodynamic limit (Erdös et al. 2007). An important feature which arises in this approach is competition between randomness and nonlinearity. As we shall see, this competition has a significant effect on the SOC problem. Under most general conditions, we can expect that, when the nonlinearity is sufficiently small, the random properties dominate, giving rise to the phenomena of Anderson localization of the excitations. That means, accordingly, that diffusion is suppressed and a wave packet that is initially localized will not spread to infinity (Anderson 1958).

So, what happens with the increasing strength of nonlinearity? The question is far from trivial, after the suggestion (largely motivated by computer simulation) (Shepelyansky 1993; Pikovsky and Shepelyansky 2008) that a weak nonlinearity can destroy localization above some level, giving rise to unlimited spreading of the wave field along the lattice, despite the existing disorder. The dynamics of the spreading has remained a matter of debate (Pikovsky and Shepelyansky 2008; Flach et al. 2009; Wang and Zhang 2009; Iomin 2010; Krivolapov et al. 2010).

In a recent investigation of NLSE with disorder, it has been theoretically found (Milovanov and Iomin 2012) that destruction of Anderson localization in the presence of nonlinearity is a critical phenomenon − likewise to a percolation transition in random lattices. Delocalization occurs spontaneously when the strength of nonlinearity goes above a certain limit. Below that limit, the field is localized similarly to the linear case. In the analysis of this section, we bring these ideas in contact with the physics of SOC and we consider a situation in which the strength of nonlinearity is determined dynamically in terms of time depending probability of site occupancy as the system fluctuates near the critical point. It is this very specific nonlinear twist with the dynamical state of the lattice, which captures the essential key signatures

[10] Essentially the same approach applies to the electrons.

due to SOC, and which as is suggested below enables to cast the SOC problem into the mathematical formalism of a discrete NLSE with disorder. The effect this twist has on the dynamics is that the localization-delocalization transition is turned self-organized, occurring exactly at the state of critical percolation.

With the aid of our DPRW lattice model, we can essentially simplify the analysis if we consider that the transport in the vicinity of the critical state occurs as a result of hole hopping between the nearest-neighbor sites, occupied by the polarization charges. As the percolation threshold is approached, we can envisage clusters of the polarization charges as the effective (random) medium, acting as a random potential on the hole wave function. Following Anderson (1958), we adopt the tight binding description for the hopping processes. Consequently, we introduce a Hamiltonian, paving the way to consider the transport problem for SOC as essentially a Hamiltonian problem. This approach poses a theoretical challenge, as it aims to connect SOC with first-principle models.

4.7.2 DANSE Equation

Focusing on the hopping motions on a lattice, we consider a variant of the discrete Anderson nonlinear Schrödinger equation (DANSE) with randomness

$$i\hbar\frac{\partial\psi_n}{\partial t} = \hat{H}_L\psi_n + \zeta|\psi_n|^2\psi_n, \tag{4.46}$$

where

$$\hat{H}_L\psi_n = \varepsilon_n\psi_n + V(\psi_{n+1} + \psi_{n-1}), \tag{4.47}$$

$\hat{H}_L$ is the Hamiltonian of the linear problem in the tight-binding approximation (Anderson 1958); ψ_n is the hole wave function; $\hat{H}_L\psi_n$ describes hopping-like transitions between the nearest-neighbor sites on a lattice; and $\zeta|\psi_n|^2\psi_n$ accounts for the generic nonlinearity of the wave process. In the above, ζ characterizes the strength of nonlinearity;[11] on-site energies ε_n are randomly distributed with zero mean across a finite energy range; V is hopping matrix element; and the total probability is normalized to $\sum_n |\psi_n|^2 = 1$. More general models can be obtained by replacing $|\psi_n|^2$ with $|\psi_n|^r$, for arbitrary $r > 0$. We do not consider such models here. In what follows, $\hbar = 1$ for simplicity. When $\zeta \to 0$, all eigenstates are exponentially localized (Anderson 1958). In the absence of randomness, DANSE (Eq. 4.46) is completely integrable. Adhering to the effective-medium description, we aim to comprehend

[11] We assume that the nonlinearity is repulsive ($\zeta > 0$), implying that it favors the spreading of the wave field. In the opposite case of attractive nonlinearity ($\zeta < 0$), solitons are typically found (Ablowitz and Segur, 1981; Zelenyi and Milovanov 2004; Krivolapov et al. 2010).

the spreading of the hole wave function under the action of nonlinear term in the limit $t \to +\infty$.

4.7.3 Coupled Nonlinear Oscillators

It is useful to expand the hole wave function using an orthogonal basis of the eigenstates, $\phi_{n,m}$, of the Anderson Hamiltonian, $\hat{H}_L$, enabling

$$\psi_n = \sum_m \sigma_m(t)\phi_{n,m} \quad (m = 1, 2, \dots). \tag{4.48}$$

"Orthogonal" means that $\sum_n \phi^*_{n,m}\phi_{n,k} = \delta_{m,k}$, where $\delta_{m,k}$ is Kronecker's delta and the star denotes complex conjugate. The total probability being equal to 1 implies that $\sum_n \psi^*_n\psi_n = \sum_m \sigma^*_m(t)\sigma_m(t) = 1$. For the nonlinear equation, Eq. (4.46), the dependence of the expansion coefficients, $\sigma_m(t)$, is found to be[12]

$$i\dot{\sigma}_k - \omega_k\sigma_k = \zeta \sum_{m_1,m_2,m_3} V_{k,m_1,m_2,m_3}\sigma_{m_1}\sigma^*_{m_2}\sigma_{m_3}, \tag{4.49}$$

where dot denotes time differentiation; ω_k $(k = 1, 2, \dots)$ are the eigenfrequencies of $\hat{H}_L$; and the amplitudes V_{k,m_1,m_2,m_3} are given by

$$V_{k,m_1,m_2,m_3} = \sum_n \phi^*_{n,k}\phi_{n,m_1}\phi^*_{n,m_2}\phi_{n,m_3}. \tag{4.50}$$

In deriving Eq. (4.49) we took into account that $\hat{H}_L\phi_{n,k} = \omega_k\phi_{n,k}$. Equation (4.49) corresponds to a system of coupled nonlinear oscillators with the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_{\rm int}, \quad \hat{H}_0 = \sum_k \omega_k\sigma^*_k\sigma_k, \tag{4.51}$$

$$\hat{H}_{\rm int} = \frac{\zeta}{2}\sum_{k,m_1,m_2,m_3} V_{k,m_1,m_2,m_3}\sigma^*_k\sigma_{m_1}\sigma^*_{m_2}\sigma_{m_3}. \tag{4.52}$$

Thus we have translated the hopping problem for the hole wave function into the interaction problem for coupled nonlinear oscillators on a lattice. In the above, $\hat{H}_0$ is the Hamiltonian of non-interacting harmonic oscillators and $\hat{H}_{\rm int}$ is the interaction Hamiltonian.[13] Each nonlinear oscillator with the Hamiltonian

$$\hat{h}_k = \omega_k\sigma^*_k\sigma_k + \frac{\zeta}{2}V_{k,k,k,k}\sigma^*_k\sigma_k\sigma^*_k\sigma_k \tag{4.53}$$

---

[12] Hint: substitute Eq. (4.48) into DANSE (Eq. 4.46), then multiply the both sides by $\phi^*_{n,k}$, and sum over n, remembering that the modes are orthogonal.

[13] We include self-interactions into $\hat{H}_{\rm int}$.

and the equation of motion

$$i\dot{\sigma}_k - \omega_k \sigma_k - \zeta V_{k,k,k,k} \sigma_k \sigma_k^* \sigma_k = 0 \tag{4.54}$$

represents one nonlinear eigenstate in the system − identified by its wave number k, unperturbed frequency ω_k, and nonlinear frequency shift $\Delta\omega_k = \zeta V_{k,k,k,k} \sigma_k \sigma_k^*$. Here, $V_{k,k,k,k}$ denote the diagonal matrix elements ($k = 1, 2, \dots$). Non-diagonal elements V_{k,m_1,m_2,m_3} characterize couplings between each four eigenstates with wave numbers k, m_1, m_2, and m_3. It is understood that the excitation of each eigenstate is nothing else than the spreading of the wave field in wave number space. Resonances occur between the eigenfrequencies ω_k and the frequencies posed by the nonlinear interaction terms. We have[14]

$$\omega_k = \omega_{m_1} - \omega_{m_2} + \omega_{m_3}. \tag{4.55}$$

When the resonances happen to overlap, a phase trajectory may occasionally switch from one resonance to another. As Chirikov (1960) realized, any overlap of resonances will introduce a random element to the dynamics along with some transport in phase space. Applying this argument to DANSE (Eq. 4.46), one sees that destruction of Anderson localization is limited to a set of resonances in a Hamiltonian system of coupled nonlinear oscillators (Eqs. 4.51 and 4.52), permitting, via a mutual overlap, and respective folding in phase space, a connected escape path to infinity.

4.7.4 Chaotic vs. Pseudochaotic Dynamics

At this point, the focus is on the topology of the random motions in phase space. We address an idealized situation first where the overlapping resonances densely fill the space. This is the familiar fully developed chaos, a regime that has been widely studied and discussed in the literature (e.g., Zaslavsky 1970; Zaslavsky and Chirikov 1972; Zaslavsky and Sagdeev 1988). A concern raised over this regime when applied to Eqs. (4.51) and (4.52) comes from the fact that it requires a diverging free energy reservoir in systems with a large number of interacting degrees of freedom such as SOC systems. Yet, developed chaos offers a simple toy-model for the transport as it corresponds with a well-understood, diffusive behavior.

A more general, as well as more intricate, situation occurs when the random motions coexist along with regular (in the Kolmogorov-Arnold-Moser, or KAM, sense) dynamics. If one takes this idea to its extreme, one ends up with the general problem of transport along separatrices of dynamical systems (Arnold 1978). This problem constitutes a fascinating nonlinear problem

[14] Conditions for nonlinear resonance are obtained by accounting for the nonlinear frequency shift.

that has as much appeal to the mathematician as to the physicist. An original important promotion of this problem to "large" (spatially extended) systems is due to Chirikov and Vecheslavov (1997).

This type of problem occurs for slow frequencies. Indeed one finds (Milovanov 2001, 2009) that resonance-overlap conditions are satisfied along the "percolating" orbits or separatrices of the random potential where the orbital periods diverge. The available phase space for the random dynamics can be very "narrow" in that case. In large systems, the set of separatrices can moreover be geometrically very complex and strongly shaped. Often it can be envisaged as a fractal network at percolation as for instance in random fields with sign-symmetry (Shklovsii and Efros, 1984; Isichenko 1992; Milovanov and Zimbardo 2000).

There is a fundamental difference between the above two transport regimes, however (chaotic vs. near-separatrix). The former regime is associated with an exponential loss of correlation permitting a Fokker-Planck description in the limit $t \to +\infty$. The latter regime when considered for large systems is associated with an algebraic loss of correlation instead (Zaslavsky 1994a,b, 2002), implying that the correlation time is infinite. There is no a conventional Fokker-Planck equation here, unless generalized to fractional derivatives (Zaslavsky 1994a,b; Metzler et al. 1998, 1999), nor the familiar Markovian property (i.e., that the dynamics are memoryless). On the contrary, there is an interesting interplay between randomness, fractality, and correlation, which is manifest in the fact that all Lyapunov exponents vanish in the thermodynamic limit, despite that the dynamics are intrinsically random (Milovanov 2009).

This situation of random non-chaotic dynamics with zero Lyapunov exponents, being in fact very general, has come to be known as "pseudochaos" (Zaslavsky 2002; Lyubomudrov et al. 2003; Zaslavsky and Edelman 2004). One may think of pseudochaos as occurring "at the edge" of stochasticity and chaos, thus separating fully developed chaos from domains with regular motions. There is a growing belief that the concept of pseudochaos offers a natural mathematical platform to obtain the fractional kinetic equations from first principles (Zaslavsky 2002).

In section 4.4.7 above it was argued that the phenomena of SOC could be described by fractional kinetics, which is a suitable and powerful formalism for long-range correlated behavior. Here, we lay more stress on this argument by proposing that SOC processes are as a matter of fact *pseudochaotic*. In this way of thinking one naturally bridges the concepts of fractional kinetics, non-Markovian transport, and SOC. Then the inherent "edge" character of pseudochaotic dynamics (Zaslavsky 2002) can be related with what one believes is the threshold nature of SOC processes (Bak et al. 1988). Support to this suggestion can be found in the results of Milovanov (2009, 2011).

4.7.5 Nearest-Neighbor Rule

This idea of "edge" behavior brings us to a model (Milovanov and Iomin 2012) where each nonlinear oscillator (Eq. 4.53) can only communicate with the rest of the wave field via a nearest-neighbor rule. Indeed this is the marginal regime yet permitting an escape path to infinity. We associate this regime with the onset of delocalization. Clearly, the number of coupling links is minimized in that case. Note that the nearest-neighbor rule guarantees that all interactions are local, this being an essential key element to SOC. When summing on the right-hand-side, the only combinations to be kept are, for the reasons of symmetry, $\sigma_k\sigma_k^*\sigma_k$ and $\sigma_{k-1}\sigma_k^*\sigma_{k+1}$. We have

$$i\dot{\sigma}_k - \omega_k\sigma_k = \zeta V_k\sigma_k\sigma_k^*\sigma_k + 2\zeta V_k^{\pm}\sigma_{k-1}\sigma_k^*\sigma_{k+1}, \tag{4.56}$$

where we have also denoted for simplicity $V_k = V_{k,k,k,k}$ and $V_k^{\pm} = V_{k,k-1,k,k+1}$. Equations (4.56) define an infinite ($k = 1, 2, \dots$) chain of coupled nonlinear oscillators where all couplings are local (nearest-neighbor-like). Then the interaction Hamiltonian in Eq. (4.52) is simplified to

$$\hat{H}_{\rm int} = \frac{\zeta}{2}\sum_k V_k\sigma_k^*\sigma_k\sigma_k^*\sigma_k + \zeta\sum_k V_k^{\pm}\sigma_k^*\sigma_{k-1}\sigma_k^*\sigma_{k+1}. \tag{4.57}$$

4.7.6 Pseudochaotic Dynamics on a Cayley Tree

We are now in position to introduce a simple lattice model for the transport. The key step is to observe that Eqs. (4.56) can be mapped on a Cayley tree where each node is connected to $c = 3$ neighbors (here, c is the coordination number). The mapping is defined as follows. A node with coordinate k represents a nonlinear eigenstate, or nonlinear oscillator with the equation of motion (4.54). There are exactly $c = 3$ branches at each node: one that we consider ingoing represents the complex amplitude σ_k^*, and the other two, the outgoing branches, represent the complex amplitudes σ_{k-1} and σ_{k+1} respectively. These settings are schematically illustrated in Fig. 4.7.

A Cayley tree being by its definition (e.g., Schroeder 1991) a hierarchical graph offers a suitable geometric model for infinite-dimensional spaces. We think of this graph as embedded into phase space of the Hamiltonian system of coupled nonlinear oscillators (Eqs. 4.51 and 4.52). In the thermodynamic limit, characterized by $\max\{k\} \to \infty$, in place of a Cayley tree, one uses the notion of a Bethe lattice.[15] Setting $\max\{k\} \to \infty$, we suppose that each node of the Bethe lattice hosts a nonlinear oscillator (Eq. 4.54). The bonds of the

[15] A Bethe lattice is an infinite version of the Cayley tree. To this end, a purist might prefer to say "bond" in place of "branch," but, once again, that's all about the terminology.

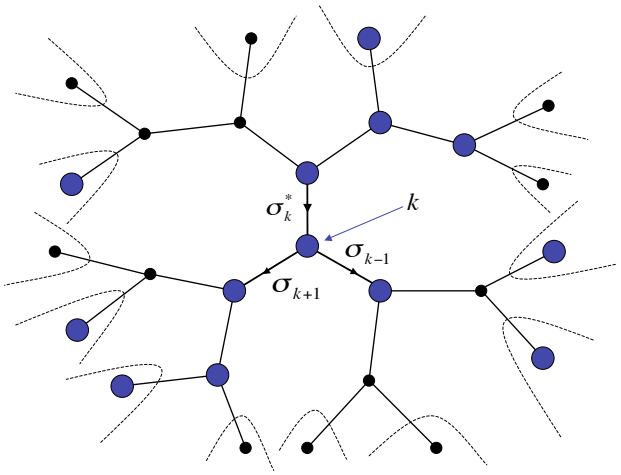


Fig. 4.7: Mapping Eqs. (4.56) on a Cayley tree. Each node represents a nonlinear eigenstate, or nonlinear oscillator with the equation of motion $i\dot{\sigma}_k - \omega_k\sigma_k - \beta V_{k,k,k,k}\sigma_k\sigma_k^*\sigma_k = 0$. Blue nodes represent oscillators in a chaotic ("dephased") state. Black nodes represent oscillators in regular state. One ingoing and two outgoing branches on node k ($k = 1, 2, \dots$) represent respectively the complex amplitudes σ_k^*, σ_{k-1}, and σ_{k+1}. Structures that are not explicitly shown are beyond the dashed lines. Adapted from Milovanov and Iomin (2012).

lattice, in their turn, can conduct oscillatory processes to their neighbors as a result of the interactions present.

Next, we assume that each oscillator can be in a chaotic ("dephased") state with the probability p (and hence, in a regular state with the probability $1 - p$). The p value being smaller than 1 implies that the domains of random motions occupy only a fraction of the lattice nodes. Whether an oscillator is dephased is decided by Chirikov's resonance-overlap condition − which may or may not be matched on node k. We believe (Chirikov and Vecheslavov 1997) that in systems with many coupled degrees of freedom each such "decision" is essentially a matter of the probability. The choice is random. Focusing on the p value, we consider system-average nonlinear frequency shift

$$\Delta\omega_{\rm NL} = \zeta\langle|\psi_n|^2\rangle_{\Delta n} \tag{4.58}$$

as an effective "temperature" of nonlinear interaction. It is this "temperature" that rules over the excitation of the various resonant "levels" in the system. With this interpretation in mind, one writes p as the Boltzmann

factor

$$p = \exp(-\delta\omega/\Delta\omega_{\rm NL}), \tag{4.59}$$

where $\delta\omega$ is the characteristic energy gap between the resonances. Expanding ψ_n over the basis of linearly localized modes, it is found that

$$\langle|\psi_n|^2\rangle_{\Delta n} = \frac{1}{\Delta n}\sum_n \sum_{m_1,m_2} \phi^*_{n,m_1}\phi_{n,m_2}\sigma^*_{m_1}\sigma_{m_2}. \tag{4.60}$$

The summation here is performed with the use of orthogonality of the basis modes. Combining with Eq. (4.58),

$$\Delta\omega_{\rm NL} = \frac{\zeta}{\Delta n}\sum_m \sigma^*_m\sigma_m. \tag{4.61}$$

The sum over m is easily seen to be equal to 1 due to the conservation of the probability. Thus, $\Delta\omega_{\rm NL} = \zeta/\Delta n$. When the field is spread over Δn states, the distance between the resonant frequencies behaves as $\delta\omega \sim 1/\Delta n$. We normalize units in Eq. (4.46) to have $\delta\omega = 1/\Delta n$ exactly. One sees that

$$p = \exp(-1/\zeta). \tag{4.62}$$

For the vanishing $\zeta \to 0$, the Boltzmann factor $p \to 0$, implying that all oscillators are in regular state. In the opposite regime of $\zeta \to \infty$, $p \to 1$. That means that all oscillators are dephased and that the random motions span the entire lattice.

There is a critical concentration, p_c, of dephased oscillators permitting an escape path to infinity for the first time. This critical concentration is nothing else than the percolation threshold on a Cayley tree. In a basic theory of percolation it is found that $p_c = 1/(c-1)$ (e.g., Schroeder 1991). This is an exact result. For $c = 3$, $p_c = 1/2$. We associate the critical value $p_c = 1/2$ with the onset of transport in the DANSE model (Eq. 4.46). When translated into the ζ values the threshold condition reads

$$\zeta_c = 1/\ln(c-1). \tag{4.63}$$

Setting $c = 3$, we have $\zeta_c = 1/\ln 2 \approx 1.4427$. This value defines the critical strength of nonlinearity that destroys the Anderson localization in the tight binding regime. For the ζ values smaller than this, the localization persists, despite that the problem is nonlinear. When $\zeta \geq 1/\ln 2$, the localization is lost and the wave field spreads to infinity.

Our conclusion so far is that destruction of Anderson localization is a thresholded phenomenon, which can be described as a percolation transition in a system of dephased oscillators on a Cayley tree (Bethe lattice). Delocalization occurs when the strength of nonlinearity, mathematically related

with the concentration of dephased oscillators (Eq. 4.62) exceeds a certain critical level. The critical point is exactly at $\zeta_c = 1/\ln 2$.

4.7.7 Making Delocalization Transition Self-Organized

With the recognition that, according to Eq. (4.62), the nonlinearity parameter, ζ, is twisted with the probability of site occupancy, p, the formalism of discrete Anderson nonlinear Schrödinger equation (Eq. 4.46), allows for a representation, which makes it possible to include the phenomena of SOC. The main idea here is to think of ζ as a fluctuating parameter, which is defined dynamically [i.e., $\zeta = \zeta(t)$] as the system evolves to percolation, and whose value is decided "on-the-fly" by the actual state of the lattice by means of $p(t) = \exp[-1/\zeta(t)]$ (see Eq. 4.62). Then the ζ value needs not to be fine tuned to its critical value, ζ_c, in order for the delocalization to occur, but it rather emerges as an attracting (singular) point as the original system of interacting charge-particles self-adjusts to percolation in response to external driving. The convergence to the critical point stems from the fact that there is a dynamical coupling between the strength of nonlinearity and fluctuating state of the lattice on which the Anderson problem is considered. Indeed the changing probability of site occupancy has feedback on the nonlinearity parameter of the model. It is this feature which leads to the dynamical rule of advancing the hole wave function and to a "self-organized" formulation of the localization-delocalization transition. We reiterate that the critical strength of nonlinearity which destroys Anderson localization is expressible in terms of the percolation threshold as $\zeta_c = -1/\ln p_c$. Generally, Eq. (4.62) shows that the behavior is *non-perturbative* in the vicinity of the criticality. Theoretically, this observation is very important as it elucidates the nature of thresholded nonlinear phenomena such as SOC. It is also manifest in the pseudochaotic character of the lattice activities, implying that the correlations persist despite that the microscopic dynamics are inherently random (Zaslavsky 2002; Zaslavsky and Edelman 2004; Milovanov 2009; Milovanov and Iomin 2012).

4.7.8 Asymptotic Spreading of the Hole Wave Function

Let us now obtain second moments for the threshold spreading of the wavefield. This task is essentially simplified if one visualizes the transport of the hole wave function as a random walk over a system of dephased oscillators. For $p \to p_c$, this system is self-similar, i.e., fractal. It is this fractal distribution of dephased oscillators which, according to Bak, Tang, and Wiesenfeld (1987), conducts "*the noise signal... through infinite distances*" just above

the marginally stable state. We consider this distribution as a percolation cluster on a Cayley tree. The fractal geometry of this cluster is fully characterized by the mean-field values of the Hausdorff dimension, $d_f = 4$, and index of anomalous diffusion, $\theta = 4$ (e.g., Nakayama et al. 1994; Havlin and ben-Avraham 2002). Note that the spectral fractal dimension is exactly 4/3 in this limit, consistenly with the original AO result (Alexander and Orbach 1982). From Eq. (4.1) one obtains (Milovanov and Iomin 2012)

$$\langle (\Delta n)^2(t) \rangle \propto t^{1/3}, \quad t \to +\infty. \tag{4.64}$$

This behavior is asymptotic in the thermodynamic limit. Note that the Hausdorff dimension being equal to 4 matches with the implication of Eqs. (4.49) and (4.50) where the coefficients V_{k,m_1,m_2,m_3} are supposed to run over 4-dimensional subsets of the ambient mapping space. Indeed it is the overlap integral of four Anderson eigenmodes (see Eq. 4.50) that decides on dimensionality of subsets of phase space where the transport processes concentrate. When the nearest-neighbor rule is applied, this overlap structure is singled out for dynamics. Under the condition that the structure is critical, i.e., "at the edge" of permitting an escape path to infinity, the support for the transport is reduced to a percolation cluster on a Bethe lattice − characterized, along with the Hausdorff dimension $d_f = 4$, by the very specific value of the connectivity exponent, $\theta = 4$. The end result is $2/(2 + \theta) = 1/3$.

We note in passing that the subdiffusion in Eq. (4.64) corresponds to an asymptotic FDE (Eq. 4.36), with the effective order of time differintegration, $\gamma = 1/3$, consistently with the single-cluster behavior in Eq. (4.1). Recently, the analysis of subdiffusion in the nonlinear Schrödinger equation with disorder by means of a fractional kinetic equation of the diffusion type has been suggested by Iomin (2010) who utilized the scheme of CTRW's. The idea was that nonlinearity-induced overlap between components of the wave-field introduces a distribution of waiting times to the hopping motion. In the above Eqs. (4.36) and (4.64) we have not as a matter of fact assumed any heavy-tailed distribution of this sort. Indeed, in our model, the random walker is supposed to take one unit step along the cluster as soon as one unit time is elapsed. Even so, with the recognition that cycles and bottlenecks of the fractal act as to delay the diffusing particle at all scales (e.g., Nakayama et al. 1994; Havlin and ben-Avraham 2002), the critical spreading of the wave-field when modeled on a regular lattice may be thought of as corresponding with the Pareto tail of an effective waiting time (w.t.) distribution, $w_\gamma(\Delta t_{\text{w.t.}}) \propto \Delta t_{\text{w.t.}}^{-4/3}$, thus sustaining the Riemann-Liouville derivative in Eq. (4.36). We find it interesting to remark that associate fractional diffusion equation (Iomin 2010) is "born" within the mathematical structure of DANSE (Eq. 4.46) containing the usual (integer) time differentiation. Indeed no *ad hoc* introduction of fractional time differentiation in the nonlinear random Schrödinger equation is needed to obtain this subdiffusion (Milovanov and Iomin 2012).

4.7.9 Summary

We have shown that the Anderson localization in disordered media can be lost in the presence of a weak nonlinearity and that the phenomenon is critical (thresholded). That means that there is a critical strength of nonlinearity above which the wave field turns to unlimited spreading. Below that limit, the field is localized similarly to the linear case. We have discussed this localization-delocalization transition as a percolation transition on the separatrix system of discrete nonlinear Schrödinger equation with disorder. This problem is solved exactly on a Bethe lattice. A threshold for delocalization is found to be $\zeta_c = 1/\ln 2 \approx 1.4427$. For the ζ values smaller than this, the localization persists, despite that the problem is nonlinear. Support for this type of behavior can be found in the results of Wang and Zhang (2009) and Krivolapov et al. (2010). More so, a "self-organized" formulation of the localization-delocalization transition has been obtained beyond perturbation theory on the basis of DANSE (Eq. 4.46) by defining the nonlinearity parameter "on-the-fly," thus utilizing its relationship (Eq. 4.62) with the probability of site occupancy. According to this picture, the occurrence of self-organized critical point relies on the dynamical feedback between the fluctuating state of the lattice and self-adjusting strength of nonlinearity. The results of these investigations when account is taken for the mathematical foundations of NLSE (Ablowitz and Segur, 1981) offer a fertile playground to describe the phenomena of SOC in connection with first principle models. In vicinity of the delocalization point the spreading of the wave field is subdiffusive, with second moments that grow with time as a powerlaw $\propto t^{1/3}$ for $t \to +\infty$. This regime bears signatures enabling to associate it with the onset of "weak" transport (Zonca et al. 2005) of Alfvén eigenmodes (AEs) in the vicinity of marginal stability of magnetic confinement systems. In this respect, we note that the characteristic aspects of sandpile physics involving SOC have, in the AE transport case, been discussed by Dendy and Helander (1997) and Chen and Zonca (2007).

4.8 The Two Faces of Nonlinearity: Instability of SOC

Often when SOC systems are said to be "nonlinear" one refers to the operation of a feedback mechanism (Kadanoff 1991; Sornette 1992a) ensuring that the control parameters need not to be fine tuned explicitly to obtain criticality. It is this feedback which stabilizes the system at the state of marginal stability, or the SOC state, as opposed to traditional critical phenomena, which do require a tuning. The comprehension of nonlinear feedback mechanism leads to another face of self-organization, the existence of a bursting instability of SOC (Milovanov 2010, 2011), which occurs in the parameter range of excessive external forcing, and for which we suggest the name "fishbone-

like" instability, by analogy with some bursting (internal-kink) instabilities in magnetic confinement systems (Chen et al. 1984). The implication is that nonlinearity can play either stabilizing as in the ideal SOC regime or destabilizing as in the regime of overdriving, role depending on the strength of interaction with the exterior. This section is aimed to discuss this topic in more detail with the aid of the DPRW model.

4.8.1 Instability Cycle

As the probability of site occupancy p approaches the percolation threshold p_c, the pair connectedness length diverges as $\xi \propto |p - p_c|^{-\nu}$. For $p > p_c$, the longest relaxation time in the system is $T_\xi \propto (p - p_c)^{-z\nu}$ and the dynamic susceptibility behaves as $\chi \propto (p-p_c)^{-z\nu\gamma}$. In the DPRW model p as a function of time is determined dynamically by the competing charge deposition and loss processes. That is, $dp/dt = Z_+ - Z_-$, where Z_+ is the net deposition rate of the free charges on the capacitor's left plate, and Z_- is the particle loss rate. The net deposition rate, or the driving rate, is the control parameter of the model: It takes a given value. The particle loss rate is obtained as electric current in the ground circuit, i.e., $Z_- = I\theta(p - p_{\min})$, where $p_{\min}$ is the lower limit of variation of p. Note that Z_- is due to the free particles leaving the system through the grounded plate. The Heaviside θ function indicates that the lattice can release charges only if/when $p \geq p_{\min}$. We expect that $p_{\min}$ lies close to, although somewhat lower than, the percolation threshold p_c. This is because the conducting cluster can still loose its charge content to the ground circuit even in the absence of a connecting path to the charged plate. The dynamics of I can be estimated from $dI/dt = \pm I/T_\xi$, where the upper sign corresponds to the relaxation process in the lattice. Putting all the various pieces together, we write

$$dp/dt = Z_+ - I\theta(p - p_{\min}), \tag{4.65}$$

$$dI/dt = WI|p - p_c|^{z\nu}\mathrm{sign}(p - p_c), \tag{4.66}$$

where $\mathrm{sign}(p - p_c) = +1$ (-1) for $p > p_c$ $(p < p_c)$ is the sign-function, and W is a numerical coefficient. Equations (4.65) and (4.66) define a simple system of equations for two cross-talking variables, the lattice occupancy per site, and the particle loss current. These model equations are perhaps the simplest nonlinear equations describing the generic fueling-storage-release cycle in driven, dissipative, thresholded dynamical systems. An examination of these equations shows that the dynamics are periodic (auto-oscillatory, see Fig. 4.8), with the peak value of electric current $I_{\max} \simeq W(p_{\max} - p_c)^{z\nu+1}$. Here, $p_{\max}$ $(p_{\max} > p_c)$ is the upper limit of the p variation. Note that $I_{\max} \to 0$ for $p_{\max} \to p_c$ as expected. The auto-oscillatory motions signify that the pure SOC state is destabilized and that the systems phase trajectories enter

the supercritical parameter range. When $p_{\max} \to 1$, the periodic dynamics acquire a sharp, bursting character. The bursts half-duration (or half-width) equals $\Delta \simeq (1/W)(p_{\max} - p_c)^{-z\nu}$. Eliminating the distance to the critical state one obtains the scaling relation $I_{\max} \propto \Delta^{-b}$, where $b = (z\nu + 1)/z\nu$. The period between the bursts is found to be $\Theta_b \simeq (p_c - p_{\min})/Z_+$. A pure SOC state with no superimposed periodic bursts arises when $\Theta_b \to \infty$. This implies that $Z_+ \to 0$ for $p \to p_c$. Thus, criticality requires the vanishing of Z_+, in agreement with the result of Vespignani and Zapperi (1998). The critical state is stable when $\Theta_b \gg T_\xi$. We have

$$(p_c - p_{\min})/Z_+ \gg T_\xi \propto |p - p_c|^{-z\nu}. \tag{4.67}$$

This is satisfied when $Z_+ \to 0$ faster than

$$Z_{+\max} \propto |p - p_c|^{z\nu}. \tag{4.68}$$

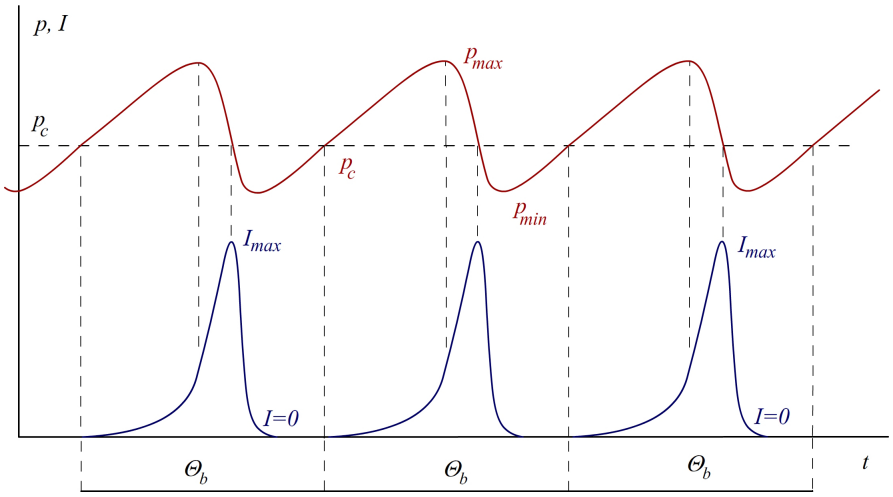


Fig. 4.8: Fishbone instability of self-organized critical state. Red: Periodic modulation in the lattice occupancy parameter, p. Blue: Periodic modulation in the particle loss current, I. The period of auto-oscillatory motion is Θ_b (much longer than the microscopic hopping time).

This limit exceeded, the system turns to auto-oscillate around the percolation point with a period dictated by the net deposition rate of the polarization charges. The physics origin of this auto-oscillatory motion lies in the fact that the changing amount of the free particles provides a feedback on the lattice occupancy parameter. It is due to this feedback relation that the DPRW system operates as a self-adjusting, intrinsically *nonlinear* dynamical system. Whether or not this feedback will excite the instability depends on

how the characteristic driving time compares to the characteristic relaxation time. Indeed, focus on the stability condition in Eq. (4.67). The system being stable at the percolation point requires that the relaxation time due to the random walks T_ξ be short compared to the characteristic driving time, $1/Z_+$. In this parameter range, any occasional charge density perturbations will dissipate via the random walks before input conducting sites are again introduced. When the percolation point is approached, because the time scale $T_\xi \propto |p - p_c|^{-z\nu}$ diverges, it is essential that the system be driven infinitesimally slowly to remain at a pure SOC state. Instability occurs when the relaxation processes operate on a longer time scale than the driving processes. In this regime, the system accumulates the polarization charges,[16] whereas to remain at criticality it would get rid of them. The accumulation of the polarization charges has a direct effect on the conductivity between the plates, which steps up with the lattice overshooting the percolation threshold. When $p_{\max} \to 1$, the system can be thought of as facing the typical conditions of electrostatic discharge in the regime of short-circuit. It should be emphasized that the feedback mechanism does a two-fold job: (i) it stabilizes the system at the state of critical percolation in a regime when the driving rate is infinitesimal; and (ii) it excites a cross-talk between the conductivity and the lattice occupancy parameters when the driving rate is faster than the relaxation rate. In the parameter range in which the strength of the driving vanishes, the multi-scale geometry of the critical percolation is dominant in providing the major transport characteristics for the DPRW lattice. The situation changes drastically when the strength of the driving increases above some level. With the systems departure away from the percolation point, the multi-scale features will soon be lost substituted by the bulk-average nonlinearities. The fact that Eqs. (4.65) and (4.66) above are formulated in terms of the system-average parameters, p and I, merely reflects that the system is allowed to appreciably depart from the state of marginal stability, or the SOC state (that means that $p_{\max}$ can be rather closer to 1 than to p_c), and that the effect of overdriving readily calls for the global features to come into play. It is noted that, in general, the multi-scale properties due to SOC can coexist along with the global or coherent features; one example of this is substorm behavior of the dynamic magnetosphere (Milovanov 2011; section 4.8.5 below).

The end result of the discussion above is that the strength of the driving plays a crucial role in dictating both linear and nonlinear behaviors in the DPRW model. To obtain a pure SOC state the driving rate should go to zero sufficiently fast as the critical point is approached. The main effect overdriving has on the DPRW dynamics is to excite unstable modes associated with periodic bursts in the particle loss current. Accordingly, the system auto-oscillates between a subcritical ($p_{\min} < p_c$) and a supercritical ($p_{\max} > p_c$) states in response to external forcing. The transition to auto-oscillatory dy-

[16] Remember that, in the proposed model, the polarization charges act as conducting states for the motion of current-carrying particles.

namics signifies the increased role of global and nonlinear behaviors in the strongly driven DPRW system as compared to a pure SOC system. The borderline between the two regimes corresponds to $T_\xi \simeq (p_c - p_{\min})/Z_+$. The stability condition $T_\xi \ll (p_c - p_{\min})/Z_+$ has serious implications for the achievable SOC regimes. It poses one important restriction on the net deposition rate of the free particles against the longest relaxation time on the incipient percolation cluster.

4.8.2 "Fishbone"-Like Instability

To help judge the result obtained, let the critical exponents take their mean-field values: $z = 2$, $\nu = 1/2$. In this limit Eqs. (4.65) and (4.66) above reproduce, up to change of variables, Eqs. (13) and (14) of Chen et al. (1984). The latter set of equations appear in a basic theory of Alfvèn instabilities as a simplified model for the coupled kink-mode and trapped-particle system in a magnetically confined toroidal plasma where beams of energetic particles are injected at high power. The mode dubbed "fishbone" is characterized by large-amplitude, periodic bursts of magnetohydrodynamic (MHD) fluctuations, which are found to correlate with significant losses of energetic beam ions (Chen et al. 1984). By comparing our Eqs. (4.65) and (4.66) with Chen et al. (1984, respective Eqs. 14 and 13 therein) one can see that the lattice occupancy per site p corresponds to the effective resonant beam-particle normalized pressure within the $q = 1$ surface (here, q is the familiar safety factor used in tokamak research); p_c corresponds to the mode excitation threshold; and the particle loss current I corresponds to the amplitude of fishbone. This direct correspondence between the two models suggests consider the instability in Eqs. (4.65) and (4.66) as analog "fishbone" instability for SOC dynamics.

This correspondence is not really surprising. Mathematically, it stems from the resonant character of the fishbone excitation, implying that the energetic particle scattering process is directly proportional to the amplitude of a fishbone (Chen et al. 1984; Chen and Zonca 2007). This resonant property dictates a specific nonlinear twist to the fishbone cycle, differentiating it from other bursting instabilities in magnetically confined plasmas. It is this "resonant" twist observed in the DPRW model system that identifies the analog "fishbone" mode for SOC. We note in passing that the existence of an instability on the top of SOC dynamics conforms with the results of Sánchez et al. (2001, 2003) in which the traditional (sandpile) SOC model has been modified by adding diffusivity, giving rise to periodic relaxation-type events

as a function of the system drive, while a pure SOC state requires a vanishing drive.[17]

4.8.3 The Threshold Character of Fishbone Excitation

The fishbone belongs to a specific class of instabilities, the energetic particle modes (EPM), which appear in a magnetic confinement system when the energetic particle pressure is comparable with the pressure of the thermal plasma (Chen 1984; Chen and Zonca 2007, 2012). The EPM constitutes a separate branch with a distinctive dispersion relation and its frequency and the growth rate crucially depend on the parameters of the energetic particle orbital motion. When the driving is strong enough, the EPM may be unstable despite having a frequency in the Alfvèn continuum where normal modes of the background plasma are typically strongly damped (the associated damping rate is proportional to the gradient of the phase velocity) (Heidbrink 2008). Instabilities in the Alfvèn continuum are often observed during intense neutral-beam injection (Chen et al. 1984; Coppi and Porcelli 1986), but they can also be excited by the energetic electrons generated experimentally by different means: electron cyclotron resonance heating (ECRH) as on DIII-D tokamak (Wong et al. 2000) and lower hybrid (LH) power injection as in Frascati Tokamak Upgrade (FTU) experiments (Zonca et al. 2007). The phenomenon is thresholded in that it requires a critical level of the absorbed power. The existence of the critical power, which we associate with the critical "driving" rate, $Z_{+\,\max}$, is well established experimentally (Fig. 4.9).

The EPM dispersion relation (Chen 1984; Chen and Zonca 2007; Chen et al. 1984; Zonca et al. 2005; reviewed in Chen and Zonca 2012) when account is taken for the well-known "fluid" (this includes the background MHD and the energetic particle adiabatic and convective responses) and "kinetic" contributions due to the energetic-particle "compressions" can be obtained via asymptotic matching procedure, leading, upon the fast and slow time scales are separated, to the frequency-dependent complex nonlinear parabolic equation, Eq. (15) of Zonca et al. (2005). A remarkable feature of this equation is that the nonlinearity due to the wave field is twisted with the free energy source term (here thought as the "driving" term). The main effect this twist has on the dynamics is that the EPMs are released in radially amplifying "avalanches" (Chen and Zonca 2007, 2012).[18] This avalanching behavior which we associate with the behavior of a strongly over-driven system in the presence of intense energetic particle population should be distinguished from the above "chain reactions of hopping motions," which are the avalanches in

[17] Beside this share, the mode referred to in Sánchez et al. (2001, 2003) is edge localized, non-resonant, fluid mode, in agreement with the diffusive nature of the added flux, but at contrast with the resonant mechanism of the fishbone excitation.

[18] To visualize, think to a large mass of mud or snow rapidly moving downhill.

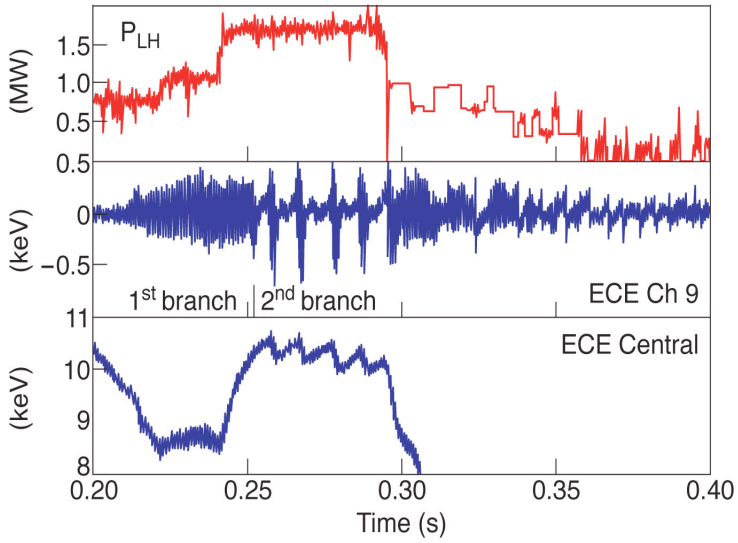


Fig. 4.9: Plots of LH coupled power, fast electron temperature fluctuations, and central radiation temperature in FTU shot # 20865. During high power LH injection, an evident transition in the electron fishbone signature takes place from almost steady state nonlinear oscillations (fixed point; marked as $1^{\rm st}$ branch) to regular bursting behavior (limit cycle; marked as $2^{\rm nd}$ branch). The transition is at $P_{\rm LH} \approx 1.69$ MW. It is noticed that the bursting behavior phase closely resembles that of well-known ion fishbones (Chen et al. 1984; Coppi and Porcelli 1986) and ECRH driven electron fishbones on DIII-D (Wong et al. 2000). Adapted from Zonca et al. 2007.

the DPRW SOC system at criticality. In the local limit, characterized by a Gaussian free energy source profile, Eq. (15) of Zonca et al. (2005) is reduced to a complex NLSE, which is different from DANSE (Eq. 4.46), in that it is dominated by the nonlinear properties, rather than by a competition with randomness.

4.8.4 Fractional Nonlinear Schrödinger Equation

If one wants to go beyond the local limit, one may use a stretched Gaussian free energy profile instead (Zonca et al. 2006). In the latter case, the resulting equation is found to be a variant of fractional nonlinear Schrödinger equation, or FNLSE (Weitzner and Zaslavsky 2003; Zelenyi and Milovanov 2004)

$$i\frac{\partial\Psi(t,x)}{\partial t} - \mathcal{D}_q\nabla^q_{-x}\nabla^q_x\Psi(t,x) + \zeta|\Psi(t,x)|^2\Psi(t,x) = 0, \tag{4.69}$$

where

$$\nabla^q_x\Psi(t,x) = [1/\Gamma(1-q)]\, x^{-q}\star'\Psi(t,x) \equiv [1/\Gamma(1-q)]\,\nabla_x\int_{-\infty}^{+\infty} dy|x-y|^{-q}\Psi(t,y) \tag{4.70}$$

is the so-called Riesz/Weyl fractional derivative (Podlubny 1999; Metzler and Klafter 2000);[19] the operator $\nabla^q_{-x}\nabla^q_x$ is a generalization of the Laplacian; x denotes respective spatial coordinate; q is a fractional exponent ($0 \le q < 1$) which corresponds with the exponent of the stretched Gaussian free energy source function (in the notation of Zonca et al. 2006, section 2, $q = \mu$); and $\mathcal{D}_q$ is a normalization constant which carries the dimension $[\mathcal{D}_q] = \mathrm{cm}^{2q}\cdot\mathrm{s}^{-1}$ (we assume that $\hbar = 1$). Mathematically, the expression

$$\nabla^{q-1}_x\Psi(t,x) = [1/\Gamma(1-q)]\, x^{-q}\star\Psi(t,x) = [1/\Gamma(1-q)]\int_{-\infty}^{+\infty} dy|x-y|^{-q}\Psi(t,y) \tag{4.71}$$

is a Fourier convolution of the wave function $\Psi(t,x)$ with a power-law ($0 \le q < 1$). One sees that the Riesz/Weyl fractional derivative is nothing else than the ordinary space derivative, ∇_x, applied to a convolution integral of the Fourier type, i.e., $\nabla^q_x\Psi(t,x) = \nabla_x\left[\nabla^{q-1}_x\Psi(t,x)\right]$. We denoted this differintegration by the symbol $\star'$ in Eq. (4.70). Although obvious, it should be emphasized that the Riesz/Weyl fractional operatior is different from the Riemann-Liouville operator in Eq. (4.35) in that the integration is performed through infinite limits. Indeed the Riemann-Liouville operator, Eq. (4.35), originates from a Laplace convolution, consistently with the initial-time problem in Eq. (4.34), as opposed to the Fourier convolution above, which has referred to integration over the configuration space in the limit $\{\max|x-y|\} \to +\infty$. Because of the improper integration in Eq. (4.71), the Riesz/Weyl operator shows a simpler behavior under transformations, as no initial values come into play (Metzler and Klafter 2000).

It is theoretically important to note that FNLSE (Eq. 4.69) is built on fractional derivatives in *space*, while time differentiation is left to be conventional (integer), likewise to DANSE (Eq. 4.46). The integer order of time differentiation in Eq. (4.69) implies that FNLSE is invariant under a one-parametric group of time translations, $t' = t + a$, making it possible to introduce a constant of motion, the energy $\mathcal{E}$. If one replaces the ordinary time differentiation here with a Riemann-Liouville derivative, Eq. (4.35), a hypothetic fractional-time FNLSE is obtained, which does not conserve the energy. On the one hand, this casts doubts on the significance of such an equation in the limit

[19] The Weyl fractional operator, ${}_{-\infty}D^q_x$, preserves the defining features of the Riesz operator, ∇^q_x, to higher ambient dimensions $d \ge 2$. See Metzler and Klafter (2000) for details. We refer to the symbol ∇^q_x as the Riesz/Weyl fractional operator as to include the higher dimensional case.

$t \to +\infty$. On the other hand, it suggests a tempting approach for the inclusion of non-stationary processes with energy amplification, as for instance convective amplification of the EPM avalanches (Zonca et al. 2005, 2006). We hasten to note, however, that we consider such non-stationary processes as being basically accomplished in the medium.

Attention we pay to the integer time differentiation in Eq. (4.69) reflects the very special role time plays in the dynamical equations originating from quantum mechanics, and in the Schrödinger equation as particular case (Iomin 2009, 2011), by contrast with the kinetic equations for transport and relaxation processes discussed above (Metzler and Klafter 2000; Metzler et al. 1998, 1999). Generally speaking, fractional time derivatives can be introduced in the quantum mechanics by means of the Wick rotation of time (Naber 2004). However, their basic physics interpretation remains to be vague: apart that they violate the Hermitian property of Hamiltonian, the corresponding notions of phase of the wave function and of semiclassical approximation have not been completely understood (Iomin 2009). On the contrary, when fractional properties are introduced, in quantum physics, by means of the Feynman propagator for Brownian path integrals (Feynman and Hibbs 1965), they naturally lead to a space fractional Schrödinger equation, which is well defined (Laskin 2000, 2002).

More generally and more importantly, the appearance of fractional derivatives in equations of motion is believed to be linked to nonlocal properties of dynamics (Zaslavsky 2002; Tarasov and Zaslavsky 2006). As the nonlocality is established, memory is introduced into the dynamical equations via coupling between respective variables in space and time. In this regard, we also note that space-fractional equations of the nonlinear Schrödinger type, including aspects of transition to chaos in a discrete NLSE with long-range interaction, have been analyzed by Weitzner and Zaslavsky (2003); Zelenyi and Milovanov (2004); Tarasov and Zaslavsky (2006); Zonca et al. (2006); and Korabel and Zaslavsky (2007). Mathematically, the fractional time Schrödinger equation was considered by Naber (2004). Its generalization to space-time fractional dynamics of nonlinear driven systems has been formulated by Chen and Zonca (2012) who associated pertinent integro-differential terms with cross-scale coupling between the many on-going wave processes interacting with an external free energy source.

In many ways the fractional NLSE (Eq. 4.69) proves to be a complementary tool in the description of nonlinear waves in complex systems, owing to its rich mathematical flavor (Zelenyi and Milovanov 2004). The asymptotics ($t \to +\infty$) of FNLSE (Eq. 4.69) are defined by a competition between nonlinearity and fractional derivative terms, which are responsible for fractional dispersion (Milovanov and Rasmussen 2002). Setting the time derivative to zero, we have $\mathcal{D}_q \nabla_{-x}^q \nabla_x^q \Psi(t,x) \sim \zeta |\Psi(t,x)|^2 \Psi(t,x)$ for $t \to +\infty$, from which the inverse power-law behavior

$$|\Psi(t,x)|^2 \sim \left[1/\Gamma(1-q)\right]^2 (\mathcal{D}_q/\zeta)|x|^{-2q} \tag{4.72}$$

can be inferred for $|x| \gg (\mathcal{D}_q/\zeta)^{1/2q}$. In the parameter range $1/2 < q < 1$, this behavior can be thought of as corresponding with the Pareto tail of a stable (Lévy) distribution of the probability density, $|\Psi(t,x)|^2$. In the latter case, the integral $\int |\Psi(t,x)|^2 dx \sim |x|^{-(2q-1)}$ converges at infinity. This convergency, other than offering normalization to the wave function, allows one to look for the wave packet-like solutions, by analogy with the integer-derivative NLSE (Ablowitz and Segur 1981). We consider a wave packet as a synchronized (coherent) state of many coupled nonlinear oscillators with long-range interaction. Of interest here are travelling synchronized states, which can be defined as nonlocal generalizations of the solitons. This type of states occurs for $\zeta > 0$. Indeed FNLSE (Eq. 4.69) contains a class of the Galilei invariant solutions that can be defined as "strange" solitary wave packets, characterized by the presence of power-law wings, Eq. (4.72), as intimated by the Riesz/Weyl fractional operator. In this class of solutions, similarly to their conventional, "integer" counterparts, the "slow" time (as opposed to the "fast" time scales of the fluctuations) is intrinsically coupled with spatial coordinate through a running-wave argument, $x-ct$. Here, c is the finite speed of the soliton. In view of nonlocality, contained in $\nabla^q_{-x}\nabla^q_x$, this dependence can work in some way as a long memory. Setting $\Psi(t,x) = \psi(x-ct)\exp(i\omega t)$, from FNLSE one arrives at the fractional envelope equation in the moving frame of the soliton

$$-\mathcal{D}_q \nabla^q_{-x}\nabla^q_x \psi(x) - \omega\psi(x) + \zeta|\psi(x)|^2\psi(x) = 0, \tag{4.73}$$

where $\omega = |\omega|$ is the frequency of synchronized oscillation, and we have assumed that $|\omega^{-1}\partial_t| \ll 1$, thus separating the "slow" and "fast" time scales. It is instructive to rewrite Eq. (4.73) in the equivalent form by allowing for the mathematical structure of a gradient on the right-hand-side:

$$-\mathcal{D}_q \nabla^q_{-x}\nabla^q_x \psi = -\frac{\partial}{\partial\psi}\left[-\frac{|\omega|}{2}\psi^2 + \frac{\zeta}{4}\psi^4\right]. \tag{4.74}$$

The latter equation describes fractional dynamics of a particle of "mass" $\mathcal{D}_q$ in the potential field $U(\psi) = -(|\omega|/2)\psi^2 + (\zeta/4)\psi^4$, where ψ represents a "coordinate," and x represents "time." The implication is that the term $-\nabla^q_{-x}\nabla^q_x\psi$ represents fractional "acceleration" along the coordinate ψ under the action of a potential force, $-\partial_\psi U(\psi) = |\omega|\psi - \zeta\psi^3$. This fractional acceleration being of non-Markovian nature will involve long-time "memory" (due to the nonlocality of the Riez/Weyl fractional differentiation with respect to x). One sees that fractional dynamics come into play via formal correspondence with associate nonlocal properties in the real space, contained in $-\nabla^q_{-x}\nabla^q_x$. This does not call for any *ad hoc* introduction of fractional time differentiation in FNLSE (Eq. 4.69), consistently with the results of the discussion above. In the local limit, holding for $q \to 1$, from Eq. (4.74) one readily finds

$$\psi(x-ct) = \sqrt{2|\omega|/\zeta}\cosh^{-1}\left[\sqrt{|\omega|/\mathcal{D}_1}(x-ct)\right]. \tag{4.75}$$

The solution of the fractional equation ($1/2 < q < 1$) is less straightforward to obtain. Somehow we expect this solution to interpolate between a stretched central core and an asymptotic inverse power-law behavior, typical for fractional-derivative equations. Indeed the asymptotic dynamics for $|x| \to +\infty$ corresponds to a particle approaching its equilibrium point at $\psi = 0$. In the vicinity of this point, the field is small, making it possible to drop the nonlinear term in Eq. (4.73). Then the solution of a linear equation, $-\mathcal{D}_q \nabla_{-x}^q \nabla_x^q \psi(x) = \omega\psi(x)$, can be obtained by standard methods, like the Fourier technique, or the method of images, yielding, in the present limit, $\psi(x-ct) \propto [|x-ct|^q \Gamma(1-q)]^{-1}$. The solution in the core region, $|x-ct| \leq (\mathcal{D}_q/\zeta)^{1/2q}$, is inferred by stretching the inverse-cosh behavior in Eq. (4.75) to fractal "time" ($\propto |x|^q$), leading to $\psi(x-ct) \propto \cosh^{-1}\left[\sqrt{|\omega|/\mathcal{D}_q}|x-ct|^q\right]$, so that the total probability, $\int_{-\infty}^{+\infty} |\psi(x)|^2 dx$, is conserved. The borderline between the two regions, $\sqrt{|\omega|/\mathcal{D}_q}|x-ct|^q \sim 1$, represents the line on which nonlinear and nonlocal properties operate on an equal footing. Note that the special cases of integer order differentiation are reinstalled through the poles of the gamma function in the limit $q \to 1$. Note, also, that the effect of stretching appears in the fact that the first derivative diverges in the origin, i.e., $\lim_{x\to\pm 0} \nabla_x \psi(x) = \mp\infty$, with the leading term in the series expansion behaving as $\propto \mp|x|^{q-1}$. Because of the sharp core, periodically occurring, nonlocal solitons can represent a bursting (fishbone-like) behavior in systems with many interacting degrees of freedom (see Figs. 4.8 and 4.9). Indeed this type of bursting behavior explains the phenomenology of fishbone near the limit cycle (2^{nd} branch in Fig. 4.9) (Zonca et al. 2007).

When cast into a more general context of the fractional parabolic equation (see Zelenyi and Milovanov 2004, Eq. 11.28 therein), FNLSE (Eq. 4.69) describes the modulation instability and associate fractional processes of self-focusing and self-compression of fracton excitations in fractal media. It also predicts an interesting phenomenon of self-delocalization (self-detrapping) of fractons (Zelenyi and Milovanov 2004). The main idea here is that the wave field can nonlinearly modify the fractal geometry of the substrate on which the excitations occur. Then it will burn off a transport "corridor" along which the Anderson localization of fractons is destroyed by nonlinearity, so that the excitations can run away from the system despite the underlying disorder. This process being essentially dissipative requires the presence of external forcing above a certain level. The present phenomenon differs from the phenomena of Anderson delocalization discussed in section 4.7 in that it is a strongly nonlinear process which is intended to significantly perturb the structural characteristics of the oscillatory medium via a nonlinear feedback reaction of the wave-field on the connectivity properties of the fractal. As a consequence, the loss of waves from the localization domain is ballistic-like, i.e., $\langle(\Delta n)^2(t)\rangle \propto t^2$, by contrast with the subdiffusion law in Eq. (4.64).

There is another version of this process, representing the EPM dynamical case discussed above (Chen 1984; Chen and Zonca 2007), in which a feed-

back occurs between the wave-field and the strength of the driving, giving rise to a ballistic loss of excitations and associate phenomena of "strong" transport in a magnetic confinement system (Zonca et al. 2005). In fact this was the version of the feedback that led to a fractional-derivative NLSE for Alfvén instabilities in the presence of intense energetic particle population (Zonca et al. 2006). Those results when account is taken for the phenomena of unstable SOC dynamics open a new perspective on the study of burning plasma as complex system (Chen and Zonca 2012; Zonca et al. 2006) beyond the restrictive assumptions of locality, vicinity to marginal stability, and the absence of correlations.

4.8.5 Mixed SOC-Coherent Behavior

The idea of "fishbone" instability in self-organized critical dynamics is very appealing as it addresses a type of behavior in which the multi-scale features due to SOC can coexist along with the global or coherent features. One example of this coexistence can be found in solar wind−magnetosphere interaction. Indeed it has been discussed by a few authors (Chang 1999; Klimas et al. 2000; Uritsky et al. 2002; Kozelov et al. 2004) that the coupled solar wind−magnetosphere−ionosphere system operates as an avalanching system and that there is a significant SOC component in the dynamics of magnetospheric storms and substorms, along with a coherent component (Sharma 1995; Chang 1999) that evolves predictably through a sequence of clearly recognizable phases (Baker et al. 1999). Here, we advocate a way of thinking (Milovanov et al. 2001b; Zelenyi and Milovanov 2004) in which the magnetospheric SOC component is associated with the properties of self-organization of electric currents and magnetic field fluctuations in the plasma sheet of the Earth's magnetotail (i.e., the "Sakura" model: section 4.6); whereas the coherent component is attributed to global instability of the cross-tail SOC current system and the phenomena of tail current disruption. The implication is that the dynamic magnetosphere survives through a *mixed* SOC-coherent behavior. In this spirit, we expect the input power due to magnetic reconnection at the Earth's dayside magnetopause to self-consistently control the departure from the state of marginal stability, or the SOC state, with stronger departures favoring the coherent features. By analogy with fishbone instability in magnetic confinement systems we suggest that the behavior is thresholded in that there is a critical input power (critical reconnection rate at the dayside magnetopause), destabilizing the SOC component in the magnetotail. At this point, a portion of the cross-tail electric current will be redirected to ionosphere − thought as analog "ground circuit" (Fig. 4.10) − leading to a decrease in the tail current intensity (tail current disruption), and to a magnetospheric disturbance, or a substorm. A model for the substorm cycle is obtained from the above coupled system of equations (Eqs. 4.65 and 4.66),

where one identifies the driving rate, Z_+, with the magnetic reconnection rate at the Earth's dayside magnetopause; the lattice occupancy parameter, p, with the average normalized energy density of magnetic fluctuation field in the magnetotail current sheet; and the particle loss current, I, with electric current in the ionosphere. We should stress that we consider a substorm as instability in the cross-tail SOC current system, which occurs on the top of self-organization to a critical state in the magnetotail, thus posing a coherent component through the dynamics.

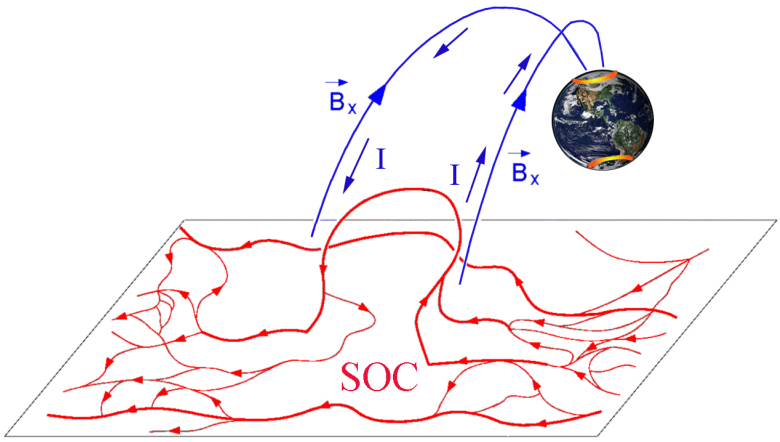


Fig. 4.10: Substorm in the Earth's magnetosphere. Red: The complex electric current system in the magnetotail current sheet. Blue: The magnetotail lobe field, $\mathbf{B}_x$. A current filament popping up in the plasma sheet interacts with the Harris-distributed magnetic field in the lobes of the magnetotail. The forces are such as to make the filament spontaneously change its orientation, favoring redirection of electric current to the ionosphere. Adapted from Milovanov (2011).

4.9 Phase Transitions in SOC Systems

It was argued that the phenomena of magnetospheric substorm bear signatures enabling to associate them with a second-order phase transition in the coupled perturbation electric current and magnetic fluctuation system (Milo-

vanov et al. 2001a,b). Indeed, when a current filament pops up in the plasma sheet, it interacts with the Harris-distributed[20] magnetic field in the lobes of the magnetotail (see Fig. 4.10). The forces are such as to make the filament spontaneously change its orientation, and in fact a local minimum in the free energy profile occurs, which favors current disruption to reinstall stability (Milovanov et al. 2001a).

Let us address the phenomena of magnetospheric substorm from a more fundamental perspective, namely, as part of the general problem of phase transitions in SOC systems. The main idea here is that some systems may spontaneously turn into a coherent state before they become SOC, since their evolution by itself drives these system to a competition between the SOC and coherent properties as a consequence of some nonlinear twist between associate order parameters. Other than substorms, this general approach may include phenomena like the L-H transition in magnetic confinement devices (Freidberg 2007), and the tokamaks as particular case, where the L-phase is associated with SOC (Carreras et al. 1998), and the H-phase is associated with spontaneously occurring coherent state.

4.9.1 Subordination to SOC

Let us consider a spatially extended system with some order parameter Υ, where the processes of self-organization develop a singularity at some value Υ_c for $t \to +\infty$. We assume that this singularity does not explicitly appear in the dynamics, implying that the system is, in this limit, critical and self-organized. The phenomena we are looking at appear when the system possesses a *competing* order parameter, which we shall denote by ψ, and for which Υ acts as input control parameter. The implication is that the dynamics of ψ is subordinated to the dynamics of Υ via some intrinsic coupling mechanism. For simplicity, we shall assume that the order parameter ψ corresponds to a coherent behavior, which we envisage as competing with the emerging multi-scale features due to SOC. Thus, while the system is developing its singular (SOC) points, it may find it thermodynamically profitable to spontaneously turn into the competing, coherent phase. As we shall see, this idea leads to a fractional extension of the Ginzburg-Landau equation (Zelenyi and Milovanov 2004; Milovanov and Rasmussen 2005), in which the conventional Laplacian is replaced by a square of fractional Riesz/Weyl derivatives.

[20] According to Harris (1962), the dependence of the lobe field is given by hyperbolic tangent of the distance to the neutral plane.

4.9.2 Generalized Free Energy Expansion

For the sake of mathematical convenience, we shall assume that the system approaches its SOC point so closely, that the geometry is, to a good approximation, self-similar (fractal). In this regime, the distribution of the competing order parameter ψ will be characterized by the diverging correlation length, owing to its coupling with Υ, so that both ψ and Υ distributions are heavy-tailed. In a sense, the fractal distribution of Υ acts a fractal support for ψ. The fact that the ψ variation involves correlations on many spatial scales must have implications for the generalized form of the free energy expansion near the phase transition point, and in particular for the gradient term (Lifshitz and Pitaevsky 1980), where the usual assumptions of locality, permitting to write this term as a simple $|\nabla_x \psi(x)|^2$, should be relaxed. Then a consistent generalization accounting for the integral effect of the correlations is obtained in terms of the Riesz/Weyl fractional operator, $\nabla_x^q \psi(x) = [1/\Gamma(1-q)]\, x^{-q} \star' \psi(x)$ (see Eq. 4.70), where q $(0 < q \leq 1)$ characterizes the strength of spatial correlation and the local limit is reinstalled for $q \to 1$. It is this convolution which we expect to replace $\nabla_x \psi(x)$ when the gradient term is considered. Indeed, the following generalized free energy expansion in vicinity of the transition point holds (Zelenyi and Milovanov 2004; Milovanov and Rasmussen 2005)

$$F = F_0 + \int_{-\infty}^{+\infty} dx \left[\mathcal{A}_q |\nabla_x^q \psi(x)|^2 + a_q |\psi(x)|^2 + \frac{1}{2} b_q |\psi(x)|^4 \right], \qquad (4.76)$$

where we have introduced three phenomenological expansion parameters $\mathcal{A}_q$, a_q, and b_q, which depend on the exponent q in general.

4.9.3 Fractional Ginzburg-Landau Equation

Varying the functional in Eq. (4.76) over the complex conjugate $\psi^*(x)$ and considering $\psi(x)$ and $\psi^*(x)$ as independent order parameters, one obtains

$$\delta F = \int_{-\infty}^{+\infty} dx \left[\mathcal{A}_q \nabla_x^q \psi(x) \nabla_x^q \delta\psi^*(x) + a_q \psi(x) \delta\psi^*(x) + b_q |\psi(x)|^2 \psi(x) \delta\psi^*(x) \right]. \qquad (4.77)$$

With use of the integration-by-parts formula (Podlubny 1999)

$$\int_{-\infty}^{+\infty} dx \varphi_1(x) \nabla_x^q \varphi_2(x) = \int_{-\infty}^{+\infty} dx \varphi_2(x) \nabla_{-x}^q \varphi_1(x), \qquad (4.78)$$

equation (4.77) implies that

$$\delta F = \int_{-\infty}^{+\infty} dx \left[\mathcal{A}_q \nabla_{-x}^q \nabla_x^q \psi(x) + a_q \psi(x) + b_q |\psi(x)|^2 \psi(x) \right] \delta\psi^*(x), \quad (4.79)$$

yielding, in view of the extremum $\delta F = 0$,

$$\mathcal{A}_q \nabla_{-x}^q \nabla_x^q \psi(x) + a_q \psi(x) + b_q |\psi(x)|^2 \psi(x) = 0. \quad (4.80)$$

Note that varying the integral in Eq. (4.76) over $\psi(x)$ leads to the conjugate equation

$$\mathcal{A}_q \nabla_{-x}^q \nabla_x^q \psi^*(x) + a_q \psi^*(x) + b_q |\psi(x)|^2 \psi^*(x) = 0, \quad (4.81)$$

which is physically identical to Eq. (4.80). Equation (4.80) has the mathematical structure of the well-known Ginzburg-Landau equation (Lifshitz and Pitaevsky 1980), in which the conventional Laplacian, ∇_x^2, is generalized to $\nabla_{-x}^q \nabla_x^q$. We consider Eq. (4.80) as the fractional Ginzburg-Landau equation, or FGLE. One sees that FGLE appears as a natural tool in describing phase transitions in SOC systems, in much the same way as the conventional Ginzburg-Landau equation describes type II phase transitions in simpler systems. Note that FGLE (Eq. 4.80) is different from the fractional envelope equation (Eq. 4.73) in that it contains the opposite sign in front of $\nabla_{-x}^q \nabla_x^q$, as well as some degree of freedom in the a_q value, as we now proceed to show.

4.9.4 The q-Exponent

Likewise to traditional type II phase transitions, one may argue that a_q changes sign at the critical point and that it linearly depends on variation of the input control parameter in the vicinity of the criticality (Landau and Lifshitz 1969; Lifshitz and Pitaevsky 1980). Then the subordination condition will imply that $a_q = \alpha_q(\Upsilon - \Upsilon_c)$ for $\Upsilon \to \Upsilon_c$, with α_q a constant which does not depend on Υ. Given that the system is driven so slowly that it develops a singular point as a result of self-organization, one predicts, with the use of FGLE (Eq. 4.80), that the distribution of the order parameter ψ will be self-similar to comply with the scaling $|\psi(x)|^2 \propto [1/\Gamma(1-q)]^2 |x|^{-2q}$ (see Eq. 4.72), from which the Hausdorff dimension $d_f = d - 2q$ can be deduced. Clearly, this scaling law stems from a competition between nonlinearity and nonlocality, contained in the fractional derivative term. Focusing on the d_f value, because the order parameter ψ is intrinsically coupled to Υ, one may expect that the ψ and Υ distributions will be essentially the same and hence, when account is taken for the percolation fractal geometry of SOC, will be both characterized by the "hyperuniversal" relation, $d_f = d - \beta/\nu$ (see section 4.1.2) where β and ν are percolation critical exponents. The latter expression will be consistent with FGLE (Eq. 4.80) when $q = \beta/2\nu$. This is the desired result. It shows that there exists a phase transition-like regime, characterized by a nontrivial fractional index of differintegration in associate fractional Ginzburg-Landau

equation, and that the behavior is nonlocal in general. Using known estimates for the parameters β and ν (Stauffer 1979; Nakayama et al. 1994), it is found that $q = 5/96$ for $d = 2$; $q \approx 0.26$ for $d = 3$; and $q \approx 0.4$ for $d = 4$. One sees that the integral $\int |\psi(x)|^2 dx$ diverges at infinity, corresponding to a spatially extended distribution of the order parameter (for $d < 6$), by contrast with the soliton-like states of FNLSE descussed above. The mean-field value, holding for $d \geq 6$, is $q = 1$. Thus the mean-field case is local, as it should. Note that locality is reinstalled through the poles of the gamma function for $q \to 1$. Setting $q \to 1$ in FGLE (Eq. 4.80), one may cross-check that the conventional Ginzburg-Landau equation is readily reintroduced. For $q < 1$, the behavior is governed by an interplay between nonlinear and nonlocal terms, likewise to the EPM excitation case (see Eq. 4.69), and its correct description requires fractional extensions of respectively the Ginzburg-Landau and nonlinear Schrödinger equations consistently with the implication of the Riesz/Weyl fractional operator.

4.10 Overall Summary and Final Remarks

The concept of self-organized criticality, or SOC, proves to be a complementary tool in drawing a physical picture of the processes underlying the dynamics of systems with many coupled degrees of freedom (i.e., "complex" systems). In this chapter we have demonstrated the diversity of appropriate mathematical methods of describing such processes, including fractional equations of the diffusion, relaxation, and Ginzburg-Landau (nonlinear Schrödinger) type, generalizing their integer-derivative counterparts, as well as the formalism of discrete Anderson nonlinear Schrödinger equation (DANSE) − extending far beyond the usual scaling theories. Some connections to Hamiltonian dynamics, paving the way to first-principle models of SOC phenomena, have been also discussed. These issues make the mathematical formalism of SOC an exciting and challenging problem.

The main emphasis in the present work has been laid on percolation, recognized as a convenient and powerful framework in describing critical phenomena in complex systems. The percolation problem finds its significance in some relation with the fundamental topology (in terms of connectedness issues) and theory of fractional manifolds (Milovanov 1997; Zelenyi and Milovanov 2004). In this respect, some elements of conformal maps of fractals, along with the percolation problem on the Riemann sphere, have been addressed. It's not the complex analysis we knew: By conformally mapping infinite percolation cluster onto the Riemann sphere we could theoretically calculate the spectral fractal dimension beyond the mean-field approaches in two embedding dimensions at criticality: $d_s = 1.327 \pm 0.001$.

To deal with dynamical problems involving feedback between the various degrees of freedom such as the SOC problem we used the idea of random walks

on a self-adjusting percolation cluster. This idea is advantageous, as it makes it possible to employ the random walks in place of the usual lattice automated redistribution rules, and in this fashion to design a family of SOC models, that are more friendly from the standpoint of their analytical treatment as compared to their CA relatives. In fact, the random walks on percolation systems offer suitable analytical forms for the diffusion, charge-conduction, and the dynamic susceptibility properties, and their significance in the study of SOC phenomena can hardly be exaggerated.

More so, we proposed a simple lattice model of self-organized criticality, the DPRW model, which addresses the SOC problem as a transport problem for electric charges (free particles and holes) on a dynamical geometry of the threshold percolation. The novel concepts of this model are: (i) a theory of self-organized criticality based on the analogy with dielectric-relaxation phenomena in self-adjusting random media, and (ii) prediction of a "resonant" instability of SOC due to the nonlinearities present. The system adjusts itself to remain at the critical point via the mechanisms of hole hopping associated with the random walk-like motion of lattice defects on a self-consistently evolving percolation cluster.

With the random walk's guide to lattice dynamics we could derive fractional analogs of the diffusion and relaxation equations, demonstrating the existence of multi-scale relaxation processes and a broad distribution of durations of relaxation events. In particular, we have shown that the relaxation to SOC of a slightly supercritical state is described by the Mittag-Leffler relaxation function, Eq. (4.22) (similarly to the Cole-Cole behavior in glassy systems: see, e.g., Coffey 2004), and not by a simple exponential function as for standard relaxation. The ideal SOC state requires that the driving rate goes to zero faster than a certain scaling law as the percolation point is approached. The model belongs to the same universality class as the BTW sandpile, and should be distinguished from the DP-like SOC models.

Thinking of holes as "excitations" of the marginally stable state, we considered a transport problem for the hole wave function in the context of DANSE equation with random potential on a lattice. An important feature which arises in this approach is competition between nonlinearity and randomness. It was argued that above a certain critical strength of nonlinearity the Anderson localization of the hole wave function is destroyed and unlimited subdiffusive spreading of the wave field along the lattice occurs. This subdiffusion process is asymptotic (Milovanov and Iomin 2012). We have seen that this problem of the critical spreading was intimately related with the outstanding problem of transport along separatrices in large systems (Chirikov and Vecheslavov 1997). With the recognition that the transition to unlimited spreading could be described as a percolation transition on a Cayley tree, a "self-organized" formulation of the phenomena of localization-delocalization in the presence of nonlinearity has been proposed. The results of this investigation have demonstrated the versatility of the DANSE formalism, which

we believe to capture the essential key elements of self-organized critical phenomena, thus offering a general analytical framework for SOC.

Overdriving the DPRW system near self-organized criticality was shown to have a destabilizing effect on the SOC state. The fundamental physics of this instability consists in the following. Because of rapid accumulation of the conducting sites, the system departs from the percolation point, and its geometry nonlinearly changes from fractal-like to crystalline-like. At this point, the conductivity of the system has greatly increased. As the lattice conducts more electricity, losses increase in the ground circuit. However, because the particle loss current has feedback on the lattice occupancy parameter, a cross-talk is excited between the systems average conductivity response and the distance to the critical state. We have observed that the instability cycle is qualitatively similar to the excitation of the internal kink ("fishbone") mode in tokamaks with high-power beam injection (the lattice occupancy per site p corresponds to the effective resonant beam-particle normalized pressure within the $q = 1$ surface (here, q is the familiar safety factor, used in tokamak research); p_c corresponds to the mode excitation threshold; and the particle loss current I corresponds to the amplitude of fishbone). The instability is "resonant" in that the particle loss process is directly proportional to I. This resonant property dictates a specific nonlinear twist to the fishbone cycle, differentiating it from other bursting instabilities in magnetically confined plasmas.

The excitation of "fishbone" instability in SOC systems leads to a type of behavior in which the multi-scale features due to SOC can coexist along with the global or coherent features (i.e., mixed SOC-coherent behavior). One example of this coexistence is found in the solar wind−magnetosphere interaction. We expect the concept of mixed SOC-coherent behavior be the plausible statistical picture for thresholded, dissipative, nonlinear dynamical systems in the parameter range of nonvanishing external forcing. In this respect, we suggest that some of the "extreme" events, or system-scale responses, observed in complex natural and social systems (e.g., Albeverio et al. 2006) may, in fact, be the fishbone-like instabilities of SOC predicted by the present theory.

It has been shown that some systems may spontaneously turn into a coherent state before they become SOC, since their evolution by itself drives these system to a competition between SOC and coherent properties as a consequence of some nonlinear twist between associate order parameters. We discussed a generalized free energy expansion for a system with extended spatial degrees of freedom, in which the order parameter due to SOC acts as input control parameter for the competing coherent behavior. Based on this expansion − which has involved the Riesz/Weyl fractional operator in place of the standard (local) gradient − a fractional generalization of the well-known Ginzburg-Landau equation has been obtained variationally (Milovanov and Rasmussen 2005). With the fractional Ginzburg-Landau equation, it should be possible to describe the much observed L-H transition in magnetic confine-

ment devices (Freidberg 2007), as well as the phenomena of magnetospheric substorm and associate tail-current disruption (Ohtani et al. 1995).

Overviewing the results obtained, we believe it will be worthwhile to pursue the above considerations not only because they arise naturally in a basic theory we are considering, but also because questions of this kind might have feedback on seemingly very diverse phenomena beyond the specialized physics context of this study. In the following we illustrate this on two examples from respectively finance and climate dynamics. Here, the discussion is unavoidably more subjective and controversial.

4.10.1 Finance

Collective phenomena in finance include as partial cases the bursting of speculative bubbles, market crashes, debt contamination (now-deepening in the euro-zone), and the spreading of bankruptcy and insolvency. A simplified toy-model here might be constructed as a variant of the DPRW model discussed above, with clusters of polarization charges thought as asset market; holes thought as insolvency; and the phenomena of hole-hopping thought as debt spreading. In this context, we might predict that, when the market is deregulated (i.e., the dynamics are random walk-like), periodic financial crises are virtually unavoidable and that the period between the crashes is inversely proportional to the rate at which speculative (not absorbed by the real economy) capital flows into the market, thus "overheating" the system. In the above formula $\Theta_b \simeq (p_c - p_{\min})/Z_+$ we associate p_c with a capitalized value at the state of marginal stability; $p_{\min}$ with the market "bottom," and Z_+ with the capital inflow rate. The net result is that financially "overheated" economies may be unstable against a "fishbone"-like oscillation − manifest in a cross-talk between capitalized values (p) and the bailouts (I). Peaks and crashes due to this oscillation might be essentially very sharp (Fig. 4.8), posing substantial risks to the investors.

More so, if we associate boundary dissipation with personal gain of the market players, we might theoretically predict that the distribution of income, Δx, in a stable market economy [21] will be given by an asymptotic inverse power-law distribution $N_\gamma(\Delta x) \propto \Delta x^{-ß}$, with the exponent ß approximately equal to 1.5, likewise to the occurrence frequency energy distribution of solar flares and/or respective distribution of faulting and earthquakes (Eq. 4.42). Interestingly, this asymptotic power-law form for the income distribution, along with the very specific value of the power exponent, ß = 1.5, was intimated by the social economist Pareto in his *Cours d'Économie Politique* as early as in 1897. More recent investigations, based on a better statistics

[21] Here, "stable" refers to the relatively quiet intervals between the crashes. We assume, accordingly, that the period $\Theta_b \simeq (p_c - p_{\min})/Z_+$ is large compared with the bursts half-width, Δ, posing an upper bound on Z_+.

available, have fairly confirmed those early results (Montroll and Shlesinger 1982; references therein).

4.10.2 Climate Dynamics

Likewise to finance, many exciting questions such as the above arise in climate research and it would be of interest to study them from the more general perspective. The focus here is on existing periodic oscillations in Earth's global climate system, as for instance the El Niño / La Niña-Southern Oscillation, or ENSO, which is a two to seven year quasi-periodic climate pattern associated with the basin-scale warming (El Niño) and cooling (La Niña) of the sea surface layer across the tropical Pacific Ocean. The extremes of this oscillation − which is identifiable in the climate reconstructions since thousands of years − are blamed for the severe weather conditions affecting climate, habitats, and the economies in many regions of the world (McPhaden et al. 2006). There are studies suggesting that ENSO involves interactions extending through different time scales with various climate phenomena, such as the seasonal cycle, interseasonal oscillations, and/or decadal oscillations. A body of studies also suggest that ENSO interacts with higher-frequency processes, as for instance intraseasonal oscillations, and with the mean state and seasonal cycle of the tropical Pacific (Guilyardi 2006). The main nonlinear processes relevant to ENSO include atmospheric convection and cloud feedbacks; wind response to ocean surface temperature anomalies; zonal advection; tropical instability waves; and thermocline-surface coupling. There is a growing belief that ENSO results from the interaction of a number of nonlinear feedbacks, either amplifying or damping the associated interannual weather anomalies. A useful resource on ENSO physics is the review by Guilyardi et al. (2009), where a report on advances made in recent years can be found, along with the challenges that lie ahead, and the related scientific debate. Another general resource is an American Geophysical Union (AGU) book, *Earth's Climate: The Ocean-Atmosphere Interaction*, edited by Wang et al. (2004), presenting a summary of current observations, theories, and models of ocean-atmosphere interaction.

Understanding and predicting ENSO is of interest from both a fundamental scientific perspective and for the practical problems of natural hazard evaluation (McPhaden et al. 2006). It was argued that ENSO was a damped mode externally sustained by atmospheric random "noise" forcing (Guilyardi et al. 2009). In this regard, an assessment of dynamical rules by which the mode is excited suggests an interesting parallel with the EPM dynamical case (Zonca et al. 2006), including features of unstable SOC dynamics (Milovanov 2010, 2011). Taking this idea to its extreme limit, one might construct a simplified yet relevant toy-model, contained in Eqs. (4.65) and (4.66) above, consisting of two cross-talking variables, p and I, such that p stands for air

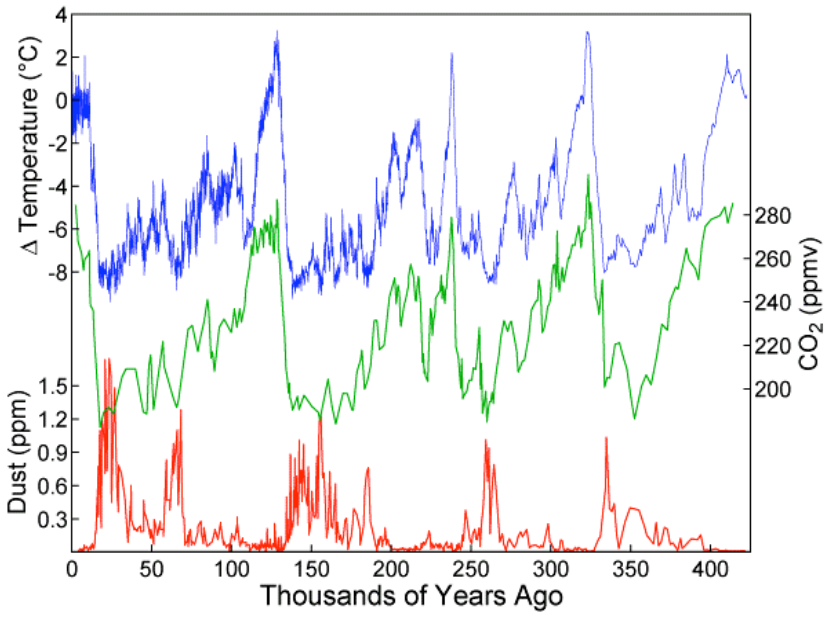


Fig. 4.11: Climate reconstruction over the past 420,000 years from the Vostok ice core, Antarctica. Blue: Temperature variation in degrees Celsius. Green: Carbon dioxide (CO_2) concentration in parts-per-million-by-volume, or p.p.m.v. Red: Dust concentration (dust concentrations are expressed in parts-per-million, or p.p.m., assuming that Antarctic dust has a density of $2,500\ \mathrm{kg{\cdot}m^{-3}}$). One sees that atmospheric concentrations of carbon dioxide correlate well with Antarctic air-temperature throughout the record, while anit-correlate with the dust concentration. Image and data credit: Petit et al. (1999).

surface pressure and I stands for ocean surface temperature. In this spirit, one might suggest that ENSO represents a climate mode of the "fishbone" type, naturally occurring in the Earth's global climate dynamics, and posing appreciable deviations from the mean state of the coupled ocean-atmosphere system. The main theoretical prediction here is that the period of the oscillation will be inversely proportional with atmospheric forcing strength. More advanced models might refer to FNLSE (Eq. 4.69) with an account for cross-scale couplings between the various wave processes involved. These theoretical investigations might open a new perspective on the introduction of fractional dynamics in climate research.

More so, the phenomenon of fishbone might shed new light on glacial-interglacial climate changes, in particular, the ∼100,000-year climate cycle,

which could be a globally induced unstable climate mode stemming from a cross-talk between air-temperature and dust concentration (see Fig. 4.11). Support for this suggestion can be found in Antarctic ice records as discussed in Petit et al. (1999). The implication is that glaciations occur naturally through functioning of Earth's climate as complex system, thus being "inherently there" as the many degrees of freedom twist and couple with the exterior. All in all, the phenomenon of fishbone warns of repeating severe events being virtually unavoidable in driven systems.

4.11 The Frontier

Generally speaking, the study of self-organized criticality phenomena is currently transitioning from an emphasis on scaling and linear-response theories to an emphasis on understanding and predicting the nonlinear dynamics of systems with many coupled degrees of freedom. In many ways these tendencies are manifest in the introduction of DANSE, FNLSE, and FGLE equations discussed above. As this transition occurs, investigations − as much experimental as theoretical and numerical − that pinpoint the nonlinear interactions in complex systems will increase in importance. Theoretically, the study of fractional-derivative, nonlinear dynamical equations is in its infancy. Beyond validation of theoretical models, the future of the field lies in the development of first-principle approaches to SOC, involving Hamiltonian approaches. These may exploit "self-organized" features of ordinary phase transitions along the lines of the DANSE and FGLE formalisms. In the geo-space plasma research, with the observations becoming multi-spacecraft and/or multi-point in scope, theoretical models are likewise to confront issues of nonlocality, self-organization, and build-up of correlations in the presence of many co-existing plasma processes (Zelenyi and Milovanov 2004; Savin et al. 2011).

Similarly to geo-space exploration, research activities in fusion plasma are now arriving at a crucial juncture that necessitates the understanding of "complexity" in the accessible and relevant operation regimes of burning plasma. Indeed it is becoming clear that the important questions that will be receiving attention in the coming years, particularly with the development of ITER and DEMO scenarios, are addressed toward the comprehension of burning plasma state as being self-organized, thresholded, nonlinear dynamical system with many interacting degrees of freedom (Chen and Zonca 2007; Zonca et al. 2005, 2006; reviewed in Chen and Zonca 2012). The ITER project is a major challenge on the way to controlled fusion burn. The specialized issues of complexity, nonlinear interactions, and SOC have found their significance in the recently formulated Fusion Advanced Studies Torus, or FAST, proposal (Pizzuto et al. 2010), promoting an European Union satellite for ITER. An emergent way of thinking here is to recognize FAST as having a

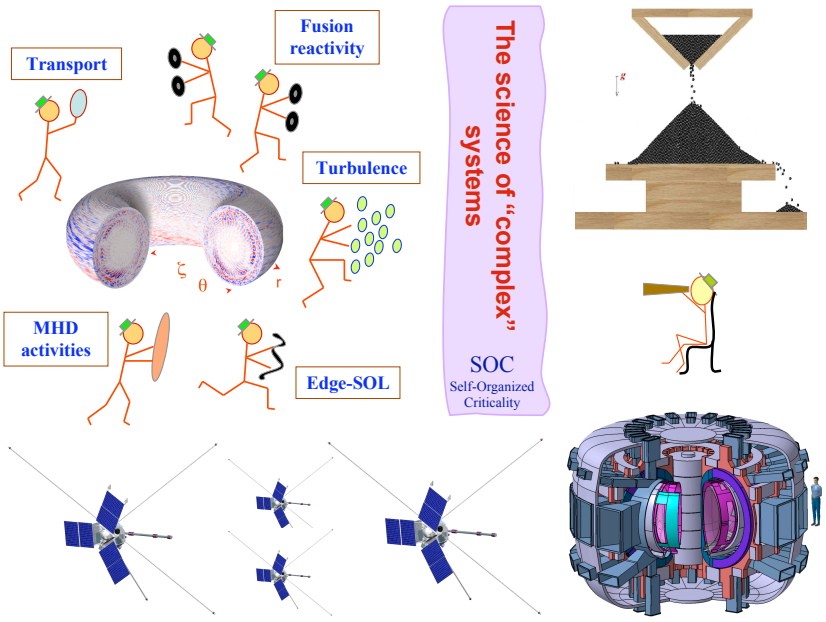


Fig. 4.12: A cartoon illustrating the many on-going plasma processes in a toroidal magnetic confinement system, viewed through the prism of complexity and self-organized criticality. A remake of Figure 1 from Zelenyi and Milovanov (2004). In the bottom-right corner is an artist's view of the FAST tokamak. Courtesy of A. Pizzuto. In the upper-right corner is a "sandpile" watch: Time is running for fusion, requiring the comprehension of SOC. The four satellites in the bottom centre represent the ROY project (Savin et al. 2011).

parallel in the geo-space exploration, the ROY mission concept (Savin et al. 2011) – a project in space research for a constellation of small, probe-like satellites,[22] aiming to investigate the dynamic magnetosphere as complex system (see Fig. 4.12). We extrapolate that the cross-disciplinary effort of bringing these exciting projects to realize will open new avenues in the study of what proves to be one of the greatest theoretical challenges in the modern nonlinear physics, the paradigm of self-organized criticality.

Acknowledgements Valuable discussions with A. Iomin, B. Kozelov, J. Juul Rasmussen, K. Rypdal, S. Savin, L. M. Zelenyi, G. Zimbardo, and F. Zonca are gratefully

[22] The Russian word "roy" means a "swarm" in English

acknowledged. I am also thankful to A. Pizzuto and F. Zonca for sharing insights with respect to the FAST project.

This work was supported by the Euratom Communities under the contract of Association between EURATOM/ENEA. Partial support was received from the ISSI project on: "Self-Organized Criticality and Turbulence."

References

Ablowitz, M.J., and Segur, H. 1981, *Solitons and the Inverse Scattering Transform*, SIAM, Philadelfia.

Albeverio, S., Jentsch, V., and Kantz, H. (Eds.) 2006, *Extreme Events in Nature and Society*, Springer, Berlin.

Alexander, S. and Orbach, R. 1982, "Density of states on fractals - fractons," J. Phys. Lett. (Paris) **43**, L625-L631.

Anderson, P.W. 1958, "Absence of diffusion in certain random lattices," Phys. Rev. **109**, 1492-1505.

Arnold, V.I. 1978, *Mathematical Methods of Classical Mechanics*, Springer, Berlin.

Aschwanden, M.J. 2011, *Self-Organized Criticality in Astrophysics. The Statistics of Nonlinear Processes in the Universe*, Springer, New York, 416p.

Aschwanden, M.J. 2012, "A statistical fractal-diffusive avalanche model of a slowly-driven self-organized criticality system," Astron. Astrophys. **539**, A2 (15 p).

Bak, P. 1996, *How Nature Works*, Springer, New York.

Bak, P. and Sneppen, K. 1993, "Punctuated equilibrium and criticality in a simple model of evolution," Phys. Rev. Lett. **71**, 4083-4086.

Bak, P., Tang, C., and Wiesenfeld, K. 1987, "Self-organized criticality: An explanation of $1/f$ noise," Phys. Rev. Lett. **59**, 381-384.

Bak, P., Tang, C., and Wiesenfeld, K. 1988, "Self-organized criticality," Phys. Rev. A **38**, 364-374.

Baker, D.N., Pulkkinen, T.I., Büchner, J. and Klimas, A.J. 1999, "Substorms: A global instability of the magnetosphere-ionosphere system," J. Geophys. Res. Space Phys. **104**, 14,601-14,611.

Bauer, T.M., Baumjohann, W., Treumann, R.A., Sckopke, N., and Luhr, H. 1995, "Low-frequency waves in the near-Earth plasma sheet," J. Geophys. Res. Space Phys. **100**, 9605-9617.

ben-Avraham, D. and Havlin, S. 2000, *Diffusion and Reactions in Fractals and Disordered Systems*, Cambridge University Press, Cambridge.

Bouchaud, J.-P. and Georges, A. 1990, "Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications," Phys. Rep. **195**, 127-293.

Brissaud, A. and Frisch, U. 1974, "Solving linear stochastic differential equations," J. Math. Phys. **15**, 524-534.

Broadbent, S.R. and Hammerslay, J.M. 1957, "Percolation processes, I. Crystals and mazes," Proc. Cambridge Philos. Soc. **53**, 629-641.

Bruggeman, D.A.G. 1935, "Berechnung verschiedener physikalischer Konstanten von heterogenen Substanzen. I. Dielektrizitätskonstanten und Leitfähigkeiten der Mischkörper aus isotropen Substanzen," Ann. Phys. (Leipzig) **24**, 636-679.

Bunde, A., Kropp, J., and Schellnhuber, H.J. (Eds.) 2002, *The Science of Disasters. Climate Disruptions, Heart Attacks, and Market Crashes*, Springer, Berlin.

Capaccioli, S., Lucchesi, M., Rolla, P.A., and Ruggeri, G. 1998, "Dielectric response analysis of a conducting polymer dominated by the hopping charge transport," J. Phys.: Condens. Matter **10**, 5595-5617.

Carreras, B.A., van Milligen, B., Pedrosa, M.A., Balbín, R., Hidalgo, C., Newman, D.E., Sánchez, E., Frances, M., García-Cortés, I., Bleuel, J., Endler, M., Davies, S., and Matthews, G.F. 1998, "Long-range time correlations in plasma edge turbulence," Phys. Rev. Lett. **80**, 4438-4441.

Carreras, B.A., van Milligen, B., Hidalgo, C., Balbin, R., Sanchez, E., Garcia-Cortes, I., Pedrosa, M.A., Bleuel, J., and Endler, M. 1999, "Self-similarity properties of the probability distribution function of turbulence-induced particle fluxes at the plasma edge," Phys. Rev. Lett. **83**, 3653-3556.

Carreras, B.A., Lynch, V.E., and Zaslavsky, G.M. 2001, "Anomalous diffusion and exit time distribution of particle tracers in plasma turbulence model," Phys. Plasmas **8**, 5096-5103.

Chang, T. 1999, "Self-organized criticality, multi-fractal spectra, sporadic localized reconnections and intermittent turbulence in the magnetotail," Phys. Plasmas **6**, 4137-4145.

Charbonneau, P., McIntosh, S.W., Liu, H.-L., and Bogdan, T.J., 2001, "Avalanche models for solar flares," Sol. Phys. **203**, 321-353.

Chen, L. 1984, "Theory of magnetohydrodynamic instabilities excited by energetic particles in tokamaks," Phys. Plasmas **1** (5), 1519-1522.

Chen, L. and Zonca, F. 2007, "Theory of Alfvèn waves and energetic particle physics in burning plasmas," Nucl. Fusion **47**, S727-S734.

Chen, L. and Zonca, F. 2012, "Physics of Alfvén waves and energetic particles in burning plasmas," Rev. Mod. Phys. (to be submitted).

Chen, L., White, R.B., and Rosenbluth, M.N. 1984, "Excitation of internal kink modes by trapped energetic beam ions," Phys. Rev. Lett. **52**, 1122-1125.

Chirikov, B.V. 1960, "Resonance processes in magnetic traps," J. Nucl. Energy, Part C: Plasma Phys. **1**, 253-260.

Chirikov, B.V. and Vecheslavov, V.V. 1997, "Arnold diffusion in large systems," Zh. Éksp. Teor. Fiz. **112**, 1132-1146.

Christon, S.P., Williams, D.J., Mitchell, D.G., Frank, L.A., and Huang, C.Y. 1989, "Spectral characteristics of plasma sheet ion and electron populations during undisturbed geomagnetic conditions," J. Geophys. Res. Space Phys. **94**, 13,409-13,424.

Coffey, W.T. 2004, "Dielectric relaxation: an overview," J. Molecular Liquids **114**, 5-25.

Coniglio, A. 1982, "Cluster structure near the percolation threshold," J. Phys. A **15**, 3829-3844.

Coppi, B., and Porcelli, F. 1986, "Theoretical model of fishbone oscillations in magnetically confined plasmas," Phys. Rev. Lett. **57**, 2272-2275.

Crosby, N.B., Aschwanden, M.J., and Dennis, B.R. 1993, "Frequency distributions and correlations of solar X-ray flare parameters," Sol. Phys. **143**, 275-299.

de Gennes, P.G. 1976, "La percolation: Un concept unificateur," La Recherche **7**, 919-927.

Dendy, R.O. and Helander, P. 1997, "Sandpiles, silos and tokamak phenomenology: a brief review," Plasma Phys. Control. Fusion **39**, 1947-1961.

Druger, S.D., Ratner, M.A., and Nitzan, A. 1985, "Generalized hopping model for frequency-dependent transport in a dynamically disordered medium, with applications to polymer solid electrolytes," Phys. Rev. B **31**, 3939-3947.

Dupree, T.H., 1967, "Nonlinear theory of drift-wave turbulence and enhanced diffusion," Phys. Fluids **10**, 1049-1056.

Dyre, J.C. and Schrøder, T.B. 2000, "Universality of ac conduction in disordered solids," Rev. Mod. Phys. **72**, 873-892.

Erdös, L., Schlein, B., and Yau, H.T. 2007, "Rigorous derivation of the Gross-Pitaevskii equation," Phys. Rev. Lett. **98**, 040404 (4p).

Feder, J. 1988, *Fractals*, Plenum, New York.

Feynman, R.P. and Hibbs, A.R. 1965, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New-York.

Flach, S., Krimer, D.O., and Skokos, Ch. 2009, "Universal spreading of wave packets in disordered nonlinear systems," Phys. Rev. Lett. **102**, 024101 (4p).

Freidberg, J. 2007, *Plasma Physics and Fusion Energy*, Cambridge Univ. Press, Cambridge (671p).

Gefen, Y., Aharony, A., and Alexander, S. 1983, "Anomalous diffusion on percolating clusters," Phys. Rev. Lett. **50**, 77-80.

Gil, L. and Sornette, D. 1996, "Landau-Ginzburg theory of self-organized criticality," Phys. Rev. Lett. **76**, 3991-3994.

Glöckle, W.G. and Nonnenmacher, T.F. 1993, "Fox function representation of non-Debye relaxation processes," J. Stat. Phys. **71**, 741-757.

Grassberger, P. and Zhang, Y.-C. 1996, "Self-organized formulation of standard percolation phenomena," Physica A **224**, 169-179.

Gruzinov, A.V., Isichenko, M.B., and Kalda, Ya.L. 1990, "2-dimensional turbulent diffusion," Zh. Eksp. Teor. Fiz. **97**, 476-488.

Guilyardi, E. 2006, "El Niño−mean state−seasonal cycle interactions in a multi-model ensemble," Climate Dyn. **26**, 329-348.

Guilyardi, E., Wittenberg, A., Fedorov, A., Collins, M., Wang, C., Capotondi, A., van Oldenborgh, G.J., and Stockdale, T. 2009, "Understanding El Niño in ocean-atmosphere general circulation models. Progress and challenges," Bull. Amer. Meteor. Soc. **90**, 325-340.

Gutenberg, B. and Richter, C.F. 1954, *Seismicity in the Earth and Related Phenomena*, Princeton Univ. Press, Princeton, New Jersey.

Harris, E.G. 1962, "On a plasma sheath separating regions of oppositely directed magnetic field," Nouvo Cimento **23**, 115-121.

Havlin, S. and ben-Avraham, D. 2002, "Diffusion in disordered media," Adv. Phys. **51**, 187-292.

Heermann, D.W. and Stauffer, D. 1981, "Phase-diagram for 3-dimensional correlated site-bond percolation," Z. Phys. B **44**, 339-344.

Heidbrink, W.W. 2008, "Basic physics of Alfvén instabilities driven by energetic particles in toroidally confined plasmas," Phys. Plasmas **15**, 055501 (15p).

Hoshino, M., Nishida, A., Yamamoto, T., and Kokubun, S. 1994, "Turbulent magnetic field in the distant magnetotail: Bottom-up process of plasmoid formation?" Geophys. Res. Lett. **21**, 2935-2938.

Hudson, H.S. 1991, "Solar flares, microflares, nanoflares, and coronal heating," Sol. Phys. **133**, 357-369.

Hurst, H.E. 1951, "Long-term storage capacity of reservoirs," Trans. Am. Soc. Civ. Eng. **116**, 770-808.

Iomin, A. 2009, "Fractional-time quantum dynamics," Phys. Rev. E **80**, 022103 (4p).

Iomin, A. 2010, "Subdiffusion in the nonlinear Schrödinger equation with disorder," Phys. Rev. E **81**, 017601 (4p).

Iomin, A. 2011, "Fractional-time Schrödinger equation: Fractional dynamics on a comb," Chaos, Solitons & Fractals **44**, 348-352.

Isichenko, M.B. 1991, "Effective plasma heat-conductivity in braided magnetic field-II. Percolation limit," Plasma Phys. Control. Fusion **33**, 809-826.

Isichenko, M.B. 1992, "Percolation, statistical topography, and transport in random media," Rev. Mod. Phys. **64**, 961-1043.

Jacobs, J.D., Koerner, H., Heinz, H., Farmer, B.L., Mirau, P., Garrett, P.H., and Vaia, R.A. 2006, "Dynamics of alkyl ammonium intercalants within organically modified montmorillonite: Dielectric relaxation and ionic conductivity," J. Phys. Chem. B **110**, 20143-20157.

Jensen, H.J. 1998, *Self-Organized Criticality. Emergent Complex Behavior in Physical and Biological Systems*, Cambridge Univ. Press, Cambridge (153p).

Kaatz, P., Prêtre, P., Meier, U., Stalder, U., Bosshard, C., Günter, P., Zysset, B., Stähelin, M., Ahlheim, M., and Lehr, F. 1996, "Relaxation processes in nonlinear optical polyimide side-chain polymers," Macromolecules **29**, 1666-1678.

Kadanoff, L.P. 1991, "Complex structures from simple systems," Phys. Today **44**, 9-11.

Kadanoff, L.P., Nagel, S.R., Wu, L., and Zhou, S.-M. 1989, "Scaling and universality in avalanches," Phys. Rev. A **39**, 6524-6537.

Klimas, A.J., Valdivia, J.A., Vassiliadis, D., Baker, D.N., Hesse, M., and Takalo, J. 2000, "Self-organized criticality in the substorm phenomenon and its relation to localized reconnection in the magnetospheric plasma sheet," J. Geophys. Res. Space Phys. **105**, 18,765-18,780.

Kohlrausch, R. 1854, "Theorie des Elektrischen Rückstandes in der Leidener Flasche," Pogg. Ann. Phys. Chem. **91**, 179-214.

Korabel, N. and Zaslavsky, G.M. 2007, "Transition to chaos in discrete nonlinear Schrödinger equation with long-range interaction," Physica A **378**, 223-237.

Kozelov, B.V., Uritsky, V.M., and Klimas, A.J. 2004, "Power law probability distributions of multiscale auroral dynamics from ground-based TV observations," Geophys. Res. Lett. 31, L20804 (4p).

Krivolapov, Y., Fishman, S., and Soffer, A. 2010, "A numerical and symbolical approximation of the nonlinear Anderson model," New J. Phys. **12**, 063035 (22p).

Kubo, R. 1963, "Stochastic Liouville equations," J. Math. Phys. **4**, 174-183.

Landau, L.D. and Lifshitz, E.M. 1969, *Statistical Physics*, Pergamon Press, Oxford.

Laskin, N. 2000, "Fractals and quantum mechanics," Chaos **10**, 780-790.

Laskin, N. 2002, "Fractional Schrödinger equation," Phys. Rev. E **66**, 056108 (7p).

Lax, M. 1958, Phys. Rev. "Generalized mobility theory," **109**, 1921-1926.

Lifshitz, E.M. and Pitaevsky, L.P. 1980, *Statistical Physics. Pt. 2. Theory of Condensed Matter*, Pergamon Press, Oxford.

Litvinenko, Yu.E. 1998, "Analytical results in a cellular automation model of solar flare occurrence," Astron. Astrophys. **339**, L57-L60.

Lui, A.T.Y., Chang, C.-L., and Yoon, P.H. 1995, "Preliminary nonlocal analysis of cross-field current instability for substorm expansion onset," J. Geophys. Res. Space Phys. **100**, 19,147-19,154.

Lyubomudrov, O., Edelman, M., and Zaslavsky, G.M. 2003, "Pseudochaotic systems and their fractional kinetics," Intl. J. Mod. Phys. B **17**, 4149-4167.

Main, I. 1996, "Statistical physics, seismogenesis, and seismic hazard," Rev. Geophys. **34**, 433-462.

Mandelbrot, B.B. 1982, *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco.

Mandelbrot, B.B. and Van Ness, J. 1968, "Fractional Brownian motions, fractional noises and applications," SIAM Rev. **10**, 422-437.

Maslov, S. and Zhang, Y.-C. 1996, "Self-organized critical directed percolation," Physica A **223**, 1-6.

McPhaden, M.J., Zebiak, S.E., and Glantz, M.H. 2006, "ENSO as an integrating concept in earth science," Science **314**, 1739-1745.

Metzler, R. and Klafter, J. 2000, "The random walk's guide to anomalous diffusion: a fractional dynamics approach," Phys. Rep. **339**, 1-77.

Metzler, R. and Klafter, J. 2004, "The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics," J. Phys. A: Math. Gen. **37**, R161-R208.

Metzler, R., Klafter, J., and Sokolov, I.M. 1998, "Anomalous transport in external fields: Continuous time random walks and fractional diffusion equations extended," Phys. Rev. E **58**, 1621-1633.

Metzler, R., Barkai, E., and Klafter, J. 1999, "Deriving fractional Fokker-Planck equations from a generalised master equation," Europhys. Lett. **46**, 431-436.

Milovanov, A.V. 1997, "Topological proof for the Alexander-Orbach conjecture," Phys. Rev. E **56**, 2437-2446.

Milovanov, A.V. 2001, "Stochastic dynamics from the fractional Fokker-Planck-Kolmogorov equation: Large-scale behavior of the turbulent transport coefficient," Phys. Rev. E **63**, 047301 (4p).

Milovanov, A.V. 2009, "Pseudochaos and low-frequency percolation scaling for turbulent diffusion in magnetized plasma," Phys. Rev. E **79**, 046403 (10p).

Milovanov, A.V. 2010, "Self-organized criticality with a fishbone-like instability cycle," Europhys. Lett. **89**, 60004 (6p).

Milovanov, A.V. 2011, "Dynamic polarization random walk model and fishbone-like instability for self-organized critical systems," New J. Phys. **13**, 043034 (22p).

Milovanov, A.V. and Iomin, A. 2012, "Localization-delocalization transition on a separatrix system of nonlinear Schrödinger equation with disorder," Europhys. Lett. **100**, 10006 (6p).

Milovanov, A.V. and Rasmussen, J.J. 2001, "Critical conducting networks in disordered solids: ac universality from topological arguments," Phys. Rev. B **64**, 212203 (4p).

Milovanov, A.V. and Rasmussen, J.J. 2002, "Fracton pairing mechanism for unconventional superconductors: Self-assembling organic polymers and copper-oxide compounds," Phys. Rev. B **66**, 134505 (11p).

Milovanov, A.V., and Rasmussen, J.J. 2005, "Fractional generalization of the Ginzburg-Landau equation: an unconventional approach to critical phenomena in complex media," Phys. Lett. A **337**, 75-80.

Milovanov, A.V. and Zelenyi, L.M. 2000, "Functional background of the Tsallis entropy: "coarse-grained" systems and "kappa" distribution functions," Nonl. Proc. Geophys. **7**, 211-221.

Milovanov, A.V., and Zelenyi, L.M. 2001, "Strange Fermi processes and power-law nonthermal tails from a self-consistent fractional kinetic equation," Phys. Rev. E **64**, 052101 (4p).

Milovanov, A.V. and Zelenyi, L.M. 2002, "Nonequilibrium stationary states in the Earth's magnetotail: Stochastic acceleration processes and nonthermal distribution functions," Adv. Space Res. **30**, 2667-2674.

Milovanov, A.V., and Zimbardo, G. 2000, "Percolation in sign-symmetric random fields: Topological aspects and numerical modeling," Phys. Rev. E **62**, 250-260.

Milovanov, A.V., Zelenyi, L.M., and Zimbardo, G. 1996, "Fractal structures and power-law spectra in the distant Earth's magnetotail," J. Geophys. Res. Space Phys. **101**, 19,903-19,910.

Milovanov, A.V., Zelenyi, L.M., Zimbardo, G., and Veltri, P. 2001a, "Self-organized branching of magnetotail current systems near the percolation threshold," J. Geophys. Res. Space Phys. **106**, 6291-6307.

Milovanov, A.V., Zelenyi, L.M., Veltri, P., Zimbardo, G., and Taktakishvili, A.L. 2001b, "Geometric description of the magnetic field and plasma coupling in the near-Earth stretched tail prior to a substorm," J. Atmos. Sol.-Terr. Phys. **63**, 705-721.

Milovanov, A.V., Rypdal, K., and Rasmussen, J.J. 2007, "Stretched exponential relaxation and ac universality in disordered dielectrics," Phys. Rev. B **76**, 104201 (8p).

Milovanov, A.V., Rasmussen, J.J., and Rypdal, K. 2008, "Stretched-exponential decay functions from a self-consistent model of dielectric relaxation," Phys. Lett. A **372**, 2148-2154.

Mittag-Leffler, G. 1905, "Sur la representation analytique d'une branche uniforme d'une fonction monogene," Acta Mathematica **29**, 101-181.

Montroll, E.W. and Bendler, J.T. 1984, "On Levy (or stable) distributions and the Williams-Watts model of dielectric relaxation," J. Stat. Phys. **34**, 129-162.

Montroll, E.W. and Shlesinger, M.F. 1982, "On $1/f$ noise and other distributions with long tails," Proc. Natl. Acad. Sci. USA **79**, 3380-3383.

Montroll, E.W. and Weiss, G.H. 1965, "Random walks on lattices. II," J. Math. Phys. **6**, 167-181.

Naber, M. 2004, "Time fractional Schrödinger equation," J. Math. Phys. **45**, 3339-3352.

Nakayama, T., Yakubo, K., and Orbach, R.L. 1994, "Dynamical properties of fractal networks: Scaling, numerical simulations, and physical realizations," Rev. Mod. Phys. **66**, 381-443.

Nishida, A., Yamamoto, T., Tsuruda, K., Hayakawa, H., Matsuoka, A., Kokubun, S., Nakamura, M., and Maezawa, K. 1994, "Structure of the neutral sheet in the distant tail (x=-210 R_e) in geomagnetically quiet times," Geophys. Res. Lett. **21**, 2951-2154.

Norman, J.P., Charbonneau, P., McIntosh, S.W., and Liu, H.-L. 2001, "Waiting-time distributions in lattice models of solar flares," Astrophys. J. **557**, 891-896.

Ohtani, S., Higuchi, T., Lui, A.T.Y., and Takahashi, K. 1995, "Magnetic fluctuations associated with tail current disruption: Fractal analysis," J. Geophys. Res. Space Phys. **100**, 19,135-19,145.

Orbach, R. 1989, "Fracton dynamics," Physica D **38**, 266-272.

O'Shaughnessy, B. and Procaccia, I. 1985, "Analytical solutions for diffusion on fractal objects," Phys. Rev. Lett. **54**, 455-458.

Paczuski, M., Maslov, S., and Bak, P. 1996, "Avalanche dynamics in evolution, growth, and depinning models," Phys. Rev. E **53**, 414-443.

Pareto, V. 1897, *Cours d'Économie Politique*, Lausanne.

Pedrosa, M.A., Hidalgo, C., Carreras, B.A., Balbín, R., García-Cortés, I., Newman, D., van Milligen, B., Sánchez, E., Bleuel, J., Endler, M., Davies, S., and Matthews, G.F. 1999, "Empirical similarity of frequency spectra of the edge-plasma fluctuations in toroidal magnetic-confinement systems," Phys. Rev. Lett. **82**, 3621-3624.

Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J.-M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, V.M., Legrand, M., Lipenkov, V.Y., Lorius, C., Pépin, L., Ritz, C., Saltzman, E., and Stievenard, M. 1999, "Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica," Nature **399**, 429-436.

Phillips, J.C. 1996, "Stretched exponential relaxation in molecular and electronic glasses," Rep. Prog. Phys. **59**, 1133-1207.

Pikovsky, A.S. and Shepelyansky, D.L. 2008, "Destruction of Anderson localization by a weak nonlinearity," Phys. Rev. Lett. **100**, 094101 (4p).

Pizzuto, A., Gnesotto, F., Lontano, M., Albanese, R., Ambrosino, G., Apicella, M.L., Baruzzo, M., Bruschi, A., Calabrò, G., Cardinali, A., Cesario, R., Crisanti, F., Cocilovo, V., Coletti, A., Coletti, R., Costa, P., Briguglio, S., Frosi, P., Crescenzi, F., Coccorese, V., Cucchiaro, A., Esposito, B., Fogaccia, G., Giovannozzi, E., Granucci, G., Maddaluno, G., Maggiora, R., Marinucci, M., Marocco, D., Martin, P., Mazzitelli, G., Mirizzi, F., Nowak, S., Paccagnella, R., Panaccione, L., Ravera, G.L., Orsitto, F., Pericoli Ridolfini, V., Ramogida, G., Rita, C., Santinelli, M., Schneider, M., Tuccillo, A.A., Zag?rski, R., Valisa, M., Villari, R., Vlad, G., and Zonca, F. 2010, "The Fusion Advanced Studies Torus (FAST): a proposal for an ITER satellite facility in support of the development of fusion energy," Nucl. Fusion **50**, 095005 (16p).

Podlubny, I. 1999, *Fractional Differential Equations*, Academic Press, San Diego.

Pommois, P., Veltri, P., and Zimbardo, G. 2001, "Kubo number and magnetic field line diffusion coefficient for anisotropic magnetic turbulence," Phys. Rev. E **63**, 066405 (5p).

Rammal, R. and Toulouse, G. 1983, "Random-walks on fractal structures and percolation clusters," J. Phys. Lett. (Paris) **44**, L13-L22.

Reuss, J.-D. and Misguich, J.H. 1996, "Low-frequency percolation scaling for particle diffusion in electrostatic turbulence," Phys. Rev. E **54**, 1857-1869.

Sánchez, R., Newman, D.E., and Carreras, B.A. 2001, "Mixed SOC diffusive dynamics as a paradigm for transport in fusion devices," Nucl. Fusion **41**, 247-256.

Sánchez, R., Newman, D.E., Carreras, B.A., Woodard, R., Ferenbaugh, W., and Hicks, H.R. 2003, "Modelling of ELM-like phenomena via mixed SOC-diffusive dynamics," Nucl. Fusion **43**, 1031-1039.

Savin, S., Zelenyi, L., Amata, E., Budaev, V., Buechner, J., Blecki, J., Balikhin, M., Klimov, S., Korepanov, V.E., Kozak, L., Kudryashov, V., Kunitsyn, V., Lezhen, L., Milovanov, A.V., Nemecek, Z., Nesterov, I., Novikov, D., Panov, E., Rauch, J.L., Rothkaehl, H., Romanov, S., Safrankova, J., Skalsky, A., and Veselov, M. 2011, "ROY−A multiscale magnetospheric mission," Planet. Space Sci. **59**, 606-617.

Scher, H. and Lax, M. 1973, "Stochastic transport in a disordered solid. I. Theory," Phys. Rev. B. **7**, 4491-4502.

Schneider, W.R. and Wyss, W. 1989, "Fractional diffusion and wave equations," J. Math. Phys. **30**, 134-144.

Schroeder, M.R. 1991, *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*, Freeman, New York.

Sharma, A.S. 1995, "Assessing the magnetospheres nonlinear behavior − its dimension is low, its predictability," Rev. Geophys. **33**, 645-650.

Sharma, A.S., Zelenyi, L.M., Malova, H.V., Popov, V.Yu., and Delcourt, D.C. 2006, "Multilayered structure of thin current sheets: multiscale "Matreshka" model," Int. Conf. Substorms **8**, 279-284.

Shepelyansky, D.L. 1993, "Delocalization of quantum chaos by weak nonlinearity," Phys. Rev. Lett. **70**, 1787-1790.

Shklovskii, B.I. and Efros, A.L. 1984, *Electronic Properties of Doped Semiconductors*, Springer, New York.

Shlesinger, M.F., Zaslavsky, G.M., and Klafter, J. 1993, "Strange kinetics," Nature (London) **363**, 31-37.

Sokolov, I.M., Klafter, J., and Blumen, A. 2002, "Fractional kinetics," Phys. Today **55**, 48-54.

Sornette, D. 1992a, "Critical phase transitions made self-organized: a dynamical system feedback mechanism for self-organized criticality," J. Phys. I France **2**, 2065-2073.

Sornette, D. 1992b, "Mean-field solution of a block-spring model of earthquakes," J. Phys. I France **2**, 2089-2096.

Stanley, H.E. and Coniglio, A. 1984, "Flow in porous-media − The backbone fractal at the percolation-threshold," Phys. Rev. B **29**, 522-524.

Stauffer, D. 1979, "Scaling theory of percolation clusters," Phys. Rep. **54**, 1-74.

Stauffer, D. and Aharony, A. 1992, *Introduction to Percolation Theory*, Taylor & Francis, London.

Straley, J.P. 1980, "The ant in the labyrinth − Diffusion in random networks near the percolation-threshold," J. Phys. C **13**, 2991-3002.

Tang, C. and Bak, P. 1988, "Critical exponents and scaling relations for self-organized critical phenomena," Phys. Rev. Lett. **60**, 2347-2350.

Tarasov, V.E. and Zaslavsky, G.M. 2006, "Fractional dynamics of coupled oscillators with long-range interaction," Chaos **16**, 023110 (13p).

Tsallis, C. 1988, "Possible generalization of Boltzmann-Gibbs statistics," J. Stat. Phys. **52**, 479-487.

Turcotte, D.L. 1999, "Self-organized criticality," Rep. Prog. Phys. **62**, 1377-1429.

Uritsky, V.M., Klimas, A.J., Vassiliadis, D., Chua, D., and Parks, G. 2002, "Scale-free statistics of spatiotemporal auroral emission as depicted by Polar UVI images: Dynamic magnetosphere is an avalanching system," J. Geophys. Res. Space Phys. **107**, 1426 (11p).

Vespignani, A. and Zapperi, S. 1998, "How self-organized criticality works: A unified mean-field picture," Phys. Rev. E **57**, 6345-6362.
Vlahos, L., Georgoulis, M., Kluiving, R., and Paschos, P. 1995, "The statistical flare," Astron. Astrophys. **299**, 897-911.
Wang, C., Xie, S.-P., and Carton, J.A. (Eds.) 2004, *Earth's Climate: The Ocean-Atmosphere Interaction*, Amer. Geophys. Union, Washington, D.C.
Wang W.-M. and Zhang, Z. 2009, "Long time Anderson localization for the nonlinear random Schrödinger equation," J. Stat. Phys. **134**, 953-968.
Weitzner, H., and Zaslavsky, G.M., 2003, "Some applications of fractional equations," Commun. Nonlinear Sci. Numer. Simulation **8**, 273-281.
Wheatland, M.S., and Uchida, Y. 1999, "Frequency-energy distributions of flares and active region transient brightenings," Sol. Phys. **189**, 163-172.
Wilkinson, D. and Willemsen, J.F. 1983, "Invasion percolation: a new form of percolation theory," J. Phys. A: Math. Gen. **16**, 3365-3376.
Williams, G., in: Allen, G. (Ed.) 1989, *Comprehensive Polymer Science*, Pergamon Press, New York, p.311.
Williams G. and Watts, D.C. 1970, "Non-symmetrical dielectric relaxation behavior arising from a simple empirical decay function," Trans. Faraday Soc. **66**, 80-85.
Wolfgang, P. and Baschnagel, J. 1999, *Stochastic Processes: From Physics to Finance*, Springer, Berlin.
Wong, K.L., Chu, M.S., Luce, T.C., Petty, C.C., Politzer, P.A., Prater, R., Chen, L., Harvey, R.W., Austin, M.E., Johnson, L.C., La Haye, R.J., and Snider, R.T. 2000, "Internal kink instability during off-axis electron cyclotron current drive in the DIII-D tokamak," Phys. Rev. Lett. **85**, 996-999.
Zaslavsky, G.M. 1970, *Statistical Irreversibility in Nonlinear Systems*, Nauka, Moscow.
Zaslavsky, G.M. 1994a, "Renormalization group theory of anomalous transport in systems with Hamiltonian chaos," Chaos **4**, 25-33.
Zaslavsky, G.M. 1994b, "Fractional kinetic equation for Hamiltonian chaos," Physica D **76**, 110-122.
Zaslavsky, G.M. 2002, "Chaos, fractional kinetics, and anomalous transport," Phys. Rep. **371**, 461-580.
Zaslavsky, G.M. and Chirikov, B.V. 1972, "Stochastic instability of nonlinear oscillations," Phys. Usp. **14**, 549-568.
Zaslavsky, G.M. and Edelman, M.A. 2004, "Fractional kinetics: from pseudochaotic dynamics to Maxwell's Demon," Physica D **193**, 128-147.
Zaslavsky, G.M. and Sagdeev, R.Z. 1988, *Introduction to the Nonlinear Physics. From Pendulum to Turbulence and Chaos*, Nauka, Moscow.
Zelenyi, L.M. and Milovanov, A.V. 2004, "Fractal topology and strange kinetics: From percolation theory to problems in cosmic electrodynamics," Phys. Usp. **47**, 749-788.
Zelenyi, L.M., Milovanov, A.V., and Zimbardo, G. 1998, "Multiscale magnetic structure of the distant tail: Self-consistent fractal approach," in *New Perspectives on the Earth's Magnetotail*, Nishida, A., Baker, D.N., and Cowley, S.W.H. (Eds.) Geophys. Monogr. Ser. **105**, pp.321-339. Amer. Geophys. Union, Washington, D.C.
Zhang, Y.-C. 1989, "Scaling theory of self-organized criticality," Phys. Rev. Lett. **63**, 470-473.
Zimbardo, G., Veltri, P., and Pommois, P. 2000, "Anomalous, quasilinear, and percolative regimes for magnetic-field-line transport in axially symmetric turbulence," Phys. Rev. E **61**, 1940-1948.
Zonca, F., Briguglio, S., Chen, L., Fogaccia, G., and Vlad, G. 2005, "Transition from weak to strong energetic ion transport in burning plasmas," Nucl. Fusion **45**, 477-484.

Zonca, F., Briguglio, S., Chen, L., Fogaccia, G., Hahm, T.S., Milovanov, A.V., and Vlad, G. 2006, "Physics of burning plasmas in toroidal magnetic confinement devices," Plasma Phys. Control. Fusion **48**, B15-B28.

Zonca, F., Buratti, P., Cardinali, A., Chen, L., Dong, J.-Q., Long, Y.-X., Milovanov, A.V., Romanelli, F., Smeulders, P., Wang, L., Wang, Z.-T., Castaldo, C., Cesario, R., Giovanozzi, E., Marinucci, M., and Pericoli Ridolfini, V. 2007, "Electron fishbones: theory and experimental evidence," Nucl. Fusion **47**, 1588-1597.

Chapter 5
Criticality and Self-Organization in Branching Processes: Application to Natural Hazards

Álvaro Corral and Francesc Font-Clos

Abstract The statistics of natural catastrophes contains very counter-intuitive results. Using earthquakes as a working example, we show that the energy radiated by such events follows a power-law or Pareto distribution. This means, in theory, that the expected value of the energy does not exist (is infinite), and in practice, that the mean of a finite set of data in not representative of the full population. Also, the distribution presents scale invariance, which implies that it is not possible to define a characteristic scale for the energy. A simple model to account for this peculiar statistics is a branching process: the activation or slip of a fault segment can trigger other segments to slip, with a certain probability, and so on. Although not recognized initially by seismologists, this is a particular case of the stochastic process studied by Galton and Watson one hundred years in advance, in order to model the extinction of (prominent) families. Using the formalism of probability generating functions we will be able to derive, in an accessible way, the main properties of these models. Remarkably, a power-law distribution of energies is only recovered in a very special case, when the branching process is at the onset of attenuation and intensification, i.e., at criticality. In order to account for this fact, we introduce the self-organized critical models, in which, by means of some feedback mechanism, the critical state becomes an attractor in the evolution of such systems. Analogies with statistical physics are drawn. The bulk of the material presented here is self-contained, as only elementary probability and mathematics are needed to start to read.

Álvaro Corral
Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, E-08193 Barcelona, Spain, e-mail: acorral@crm.cat

Francesc Font-Clos
Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, E-08193 Barcelona, Spain
Departament de Matemàtiques, Universitat Autònoma de Barcelona, E-08193 Barcelona, Spain e-mail: fontclos@crm.cat

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

5.1 The Statistics of Natural Hazards

Only fools, charlatans and liars predict earthquakes
C. F. Richter

Men, and women, have always been threatened by the dangers of Earth: volcanic eruptions, tsunamis, earthquakes, hurricanes, floods, etc. Sadly, still in the 21st century our societies have not been able to get rid of such a sword of Damocles. But are natural catastrophes submitted to the caprices of the gods? Or do these disasters contain some hidden patterns or regularities? The first view has been dominant for many centuries in the history of humankind, and it has been only in recent times that a more rational perspective has started to consolidate.

5.1.1 The Gutenberg-Richter Law

One of the first laws quantifying the occurrence of a natural hazard was proposed for earthquakes by the famous seismologists Beno Gutenberg and Charles F. Richter in the 1940's, taking advantage from the recent development of the first magnitude scale by Richter himself. The Gutenberg-Richter law is quite simple: if one counts the number of earthquakes in any seismically active region of the world during a long enough period of time, one must find that for each 100 earthquakes of magnitude M greater or equal than 3 there are, approximately (on average), 10 earthquakes with $M \geq 4$, one earthquake with $M \geq 5$, and so on (Gutenberg and Richter 1944, Utsu 1999, Kanamori and Brodsky 2004). So, the vast majority of events are the smallest ones, and, fortunately, only very few of them can become catastrophic, maintaining a constant proportion between their number.

It is not possible to measure all earthquakes on our planet, but for some areas with very accurate seismic monitoring it has been found

that the Gutenberg-Richter law holds down to magnitude - 4 (Kwiatek et al. 2010); this corresponds to small rock cracks of a few centimeters in length (negative magnitudes are introduced to account for the fact that there can be earthquakes smaller than those of zero magnitude). And, more remarkably, for nanofracture experiments in the laboratory (Åström et al. 2006), the law has been verified up to magnitude below -13. The scarcity of the big events contained in the law leaves as open the question about which is its upper limit of validity.

Despite not being recognized or mentioned by Gutenberg and Ritchter in their original paper (1944), any reader with a minimum knowledge of probability and statistics will immediately realize that the Gutenberg-Richter law implies an exponential distribution of the magnitudes of earthquakes, i.e.,

$$D_M(M) \propto 10^{-bM}, \tag{5.1}$$

with $D_M(M)$ the probability density of M, the parameter b taking a value close to 1, and the symbol $\propto$ standing for proportionality (with the constant of proportionality ensuring proper normalization).

But which is the meaning of the Gutenberg-Richter law, in addition to provide an easy-to-remember relationship between the relative abundances of earthquakes? The interpretation depends, of course, on the meaning of magnitude, which we have avoided to define. In fact, there is not a unique magnitude, but several of them; second, magnitudes do not have physical dimensions (i.e., units); and third, "magnitudes reflect radiation only from subportions of the rupture, and they saturate above certain size, rather than giving a physical characterization of the entire earthquake source" (Ben-Zion 2008). More in-depth understanding comes from the energy radiated by an earthquake, which is believed to be an exponential function of its magnitude (Kanamori and Brodsky 2004), that is,

$$E \propto 10^{3M/2}, \tag{5.2}$$

with a proportionality factor close to 60 kJ (Utsu 1999); so, an increase by 1 in the magnitude implies an increase in energy by a factor $\sqrt{1000} \simeq 32$. Thus, an earthquake of magnitude 9 radiates as much energy as 1000 earthquakes of magnitude 7, or as 10^6 of magnitude 5.

One can reformulate then the Gutenberg-Richter law in terms of the energy. Indeed, the probability of an event is "independent" of the variable we use to describe it, and so,

$$D_E(E) = D_M(M)\frac{dM}{dE}, \tag{5.3}$$

with $D_E(E)$ the probability density of the energy. Using equation (5.2), we can express M as a function of E,

$$M \propto \log E, \tag{5.4}$$

and differentiate to obtain dM/dE,

$$\frac{dM}{dE} \propto \frac{1}{E}, \tag{5.5}$$

so that equation (5.3) reads:

$$D_E(E) \propto 10^{-bM}\frac{1}{E} = \left(10^{\frac{3M}{2}}\right)^{-\frac{2b}{3}}\frac{1}{E} = E^{-\frac{2b}{3}}\frac{1}{E}. \tag{5.6}$$

Summarizing, this straightforward change of variables leads to

$$D_E(E) \propto \frac{1}{E^{\alpha}}, \text{ with } \alpha = 1 + \frac{2b}{3}, \tag{5.7}$$

and this is just the so-called power-law distribution, or Pareto distribution (Newman 2005), with exponent α around 1.67 when b is close to 1. Notice from equation (5.7) that in order that $D_E(E)$ is a proper probability density function, it has to be defined above a minimum energy $E_{\min} > 0$, otherwise (if $E_{\min} = 0$), it cannot be normalized. Although the true value of $E_{\min}$ cannot be measured (it is too small), this parameter is not important as it does not influence any properties of earthquakes.

Figure 5.1 displays the probability density of the seismic moment for world-wide shallow earthquakes (Kagan 2010); this variable is assumed to be proportional to the energy, but much easier to measure accurately (Kanamori and Brodsky 2004), and so, it should also be power-law distributed, with the same exponent. The straight line in the plot is the defining characteristic of a power law in double logarithmic scale, as $\log D_E(E) = C - \alpha \log E$. A fit by maximum likelihood estimation (Clauset et al. 2009, Peters et al. 2010) yields $\alpha \simeq 1.68$.

Two important properties of power-law distributions are scale invariance (with some limitations due to the normalization condition) and divergence of the mean value (if the exponent α is below or equal to 2). These are explained in the Appendix.

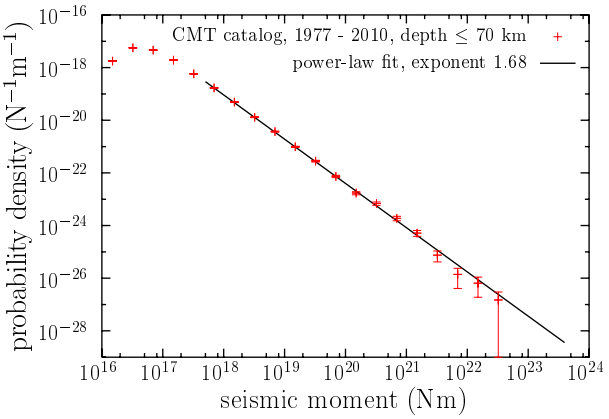


Fig. 5.1: Estimation of the probability density of seismic moment for world-wide shallow earthquakes (in log-log scale), using the so-called CMT catalog (Kagan 2010). A power law fit results in an exponent $\alpha = 1.68$. Radiated energy should give the same power law behavior. Deviations at small values of the seismic moment are attributed to the incompleteness of the catalog.

To conclude this subsection, let us mention that the power-law distribution of sizes is not a unique characteristic of earthquakes. It has been claimed that many other natural hazards are also power-law distributed, although with different exponents (and maybe with a lower or an upper cut-off): tsunamis (Burroughs and Tebbens 2005), landslides, rockfalls (Malamud 2004), volcanic eruptions (McClelland et al. 1989, Lahaie and Grasso 1998), hurricanes (Corral et al. 2010), rainfall (Peters et al. 2010), auroras (Freeman and Watkins 2002), forest fires (Malamud et al. 2005)... As the reader will figure out, some of the facts that we will explain having in mind earthquakes can also be applied to some of these natural hazards, but maybe not to all of them. It is an open question to distinguish between these different cases. For an account of power-law distributions in other areas beyond geoscience see the excellent review by Newman (2005).

5.1.2 A First Model for Earthquake Occurrence

As far as we know, a first attempt to develop an earthquake model in order to explain the Gutenberg-Richter law was undertaken by Michio Otsuka in the early 1970's (Otsuka 1971, 1972, Kanamori and Mori 2000). He used as a metaphor the popular Chinese game of go, although we will formulate the model in relation to the game of domino, probably more familiar to the potential readers.

Instead of playing domino, we are going to play a different game with their pieces. The idea is to make the domino pieces to topple, as in the well-known contests and attempts to break a Guinness world record, but with two important differences. First, the pieces are not put in a row, but, rather, they constitute a kind of tree. Second, when one piece topples, one does not know what will happen next, i.e., if some other pieces will topple in turn (and how many will) or not. So, we have a stochastic cascade process that supposedly mimics the rupture that takes place in a seismic fault during an earthquake. The tree of domino pieces constitutes the fault, and each piece is a small fault patch, or element. The earthquake is the chain reaction of toppling of pieces (i.e., failures of patches).

Getting more concrete, Otsuka assumed that the tree representing the fault had a fixed number of branches at each position, or node, and that the toppling would propagate from each branch to the next element with a fixed probability p, independently of any other variable. So, the number of propagating branches resulting from a single one would follow the binomial distribution (Ross 2002). For instance, in Fig. 5.2, the possible number of branches per element is just 2. If a fixed elementary energy is associated to the failure of each patch, one can obtain the energy released in this process from the number of topplings, allowing the comparison with the Gutenberg-Richter law, see nevertheless Sec. 4.1 of the review by Ben-Zion (2008). So, the

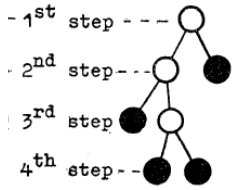


Fig. 5.2: Scheme of Otsuka's model for earthquake ruptures. White circles correspond to the propagation of the rupture, whereas black ones indicate termination points (Otsuka 1972).

propagation of ruptures is considered a probability controlled phenomenon, in such a way that when an earthquake starts, it is not possible to know how big it will become. Later, we will see that this statement is stronger than what it looks like here. The usual domino effect, in which one toppling induces a new one for sure and so on, would correspond to the controversial concept of a characteristic earthquake (Stein 2002, Ben-Zion 2008, Kagan et al. 2012), an event that always propagates along the complete fault or fault system and would release always the same amount of energy.

The novel and original model in geophysics explained in this subsection, proposed by Otsuka in the 1970's, was already known by a few mathematicians 100 years in advance. It will take us the next pages to explain the distribution of energy in this model.

5.2 Branching Processes

Besides gambling, many probabilists have been interested in reproduction
G. Grimmett and D. Stirzaker

Let us move to the Victorian (19th century) England. There, Sir Francis Galton, the polymath father of the statistical tools of correlation and regression, and cousin of Charles Darwin, was dedicated to many different affairs. In addition to the height of sons in relation to the heights of their fathers, he was concerned about the decay and even extinction of families that were important in the past, and about whether this decline was a consequence of a diminution in fertility provoked by the rise in comfort. If that were the case, population would be constantly fed by the contribution of the lower classes (Watson and Galton 1875). In order to better understand the problem, he devised a null model in which the number of sons of each men was random (the abundance of women was not considered to be a limitation). Despite the apparent simplicity of the model, Galton was not able to solve it, and made

a public call for help. The call was also fruitless, and then Galton turned to the mathematician and reverend Henry William Watson.

5.2.1 Definition of the Galton-Watson Process

Let us consider "elements" that can generate other elements and so on. These elements may represent British aristocratic men that have some male descendants, (or, in a more fresh perspective, women from anywhere that give birth to her daughters, or, perhaps more properly, bacteria that replicate), neutrons that release more neutrons in a nuclear chain reaction, or fault patches that slip during an earthquake. The Galton-Watson process assumes that each of these elements triggers a random number K of offspring elements in such a way that each K is independent from that of the other elements and all K are identically distributed, with probabilities $P(K = 0) = p_0$, $P(K = 1) = p_1$, $\ldots P(K = k) = p_k$, with $k = 0, 1, \ldots \infty$ (Harris 1963). (Naturally, the normalization condition imposes $\sum_{\forall k} p_k = 1$.)

The model starts with one single element, in what we call the zeroth generation of the process, as shown in Fig. 5.3. The K offsprings of this first element constitute the first generation. Let $Z_0 \equiv 1$ denote the number of elements of the zeroth generation, Z_1 the number of elements of the first generation, etc. Obviously, by construction, $P(Z_1 = k) = p_k$. The number of elements in the $t + 1$ generation is obtained from the number of the previous generation t as

$$Z_{t+1} = \sum_{i=1}^{Z_t} K_i, \tag{5.8}$$

with $t \geq 0$, where K_i corresponds to the number of offsprings of each element in the t generation. Equation (5.8) can be used to simulate the process in a straightforward way and will be very important to its analytical treatment, in order to calculate the probability distribution of Z_t, for any t. Some readers may recognize that the variables $Z_0, Z_1, \ldots$ form a Markov chain, but this is not relevant for our purposes. And of course, Otsuka's earthquake model is a particular case of the Galton-Watson process corresponding to a binomial distribution for $P(K = k)$.

5.2.2 Generating Functions

An extremely convenient mathematical tool will be the probability generating function (Grimmett and Stirzaker 2001). For the random variable K this is, by definition,

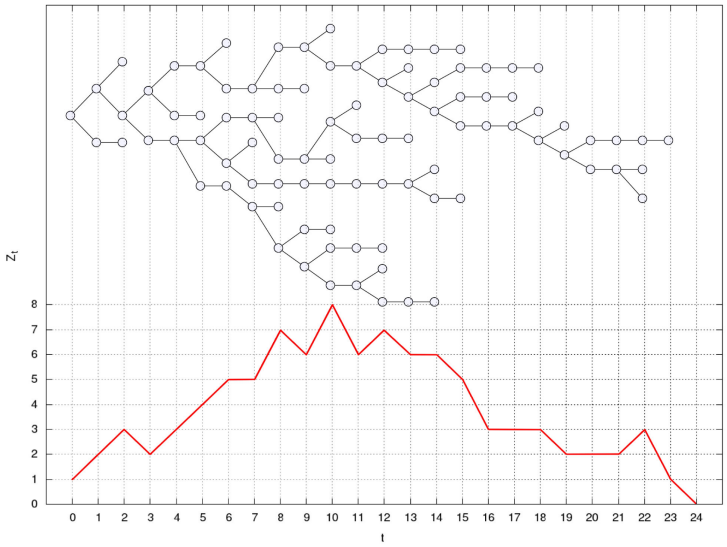


Fig. 5.3: A realization of the Galton-Watson process. At the top, the tree associated to the process is shown, starting from the left ($Z_0 = 1$). At the bottom, the evolution of the number of elements originated in each generation t are displayed. The model for $P(K = k)$ is binomial with $n = 2$ and $p = 1/2$, corresponding to the critical case (see main text).

$$f_K(x) \equiv \sum_{k=0}^{\infty} p_k x^k = \langle x^K \rangle, \tag{5.9}$$

where the brackets indicate expected value. The normalization condition guarantees that $f_K(x)$ is always defined at least in the $x-$interval $[-1, 1]$, although only the interval $[0, 1]$ will be of interest for us. Of course, the same definition applies to any other random variable; in the concrete case of K (which represents the number of offsprings of any element) we may drop the subindex, i.e., $f_K(x) = f(x)$.

Very useful and straightforward properties will be,

1. $f_K(0) = P(K = 0)$;
2. $f_K(1) = 1$ (by normalization);
3. $f'_K(1) = \sum_{\forall k} p_k k = \langle K \rangle \equiv m$;
4. $f'_K(x) \geq 0$ for $x \geq 0$ (non-decreasing function);
5. $f''_K(x) \geq 0$ for $x \geq 0$ (non-convex function, "looking from above");

the primes denoting derivatives (left-hand derivatives at $x = 1$). Note that although we illustrate these properties with the variable K, they are valid for

the generating function of any other (discrete) random variable. So, the plot of a probability generating function between 0 and 1 is very constrained. We anticipate that two main cases will exist, depending on whether the expected value of K is $m < 1$ or whether $m > 1$. This is natural, as the first case corresponds to a population that on average decreases from one generation to the next whereas in the second case the population grows, on average.

Another property but not so straightforward is that the generating function of a sum of N independent identically distributed variables K (with N fixed) is the N-th power of the generating function of K; that is, if

$$\Sigma = \sum_{i=1}^{N} K_i, \tag{5.10}$$

then

$$f_\Sigma(x) = f_K(x)^N. \tag{5.11}$$

Indeed,

$$\begin{aligned} f_\Sigma(x) = \langle x^\Sigma \rangle = \langle x^{\sum K_i} \rangle = \langle x^{K_1} \cdot x^{K_2} \cdots x^{K_N} \rangle = \\ = \langle x^{K_1} \rangle \langle x^{K_2} \rangle \cdots \langle x^{K_N} \rangle = f_K(x)^N, \end{aligned} \tag{5.12}$$

where we can factorize the expected values due to statistical independence among the K_i's.

In general, if the random variables K_i were not identically distributed (but still independent), the generating function of their sum would be the product of their generating functions. The demonstration is essentially the same as before, and one only needs to introduce new notation for the different generating functions.

A following step is to consider that N is also a random variable, with generating function $f_N(x)$. Then,

$$f_\Sigma(x) = f_N(f_K(x)). \tag{5.13}$$

Note that equation (5.13) is just a generalization of equation (5.11), i.e., now we calculate the expected value of the powers of $f_K(x)$ depending on the values that N make take. In any case, it is easy to demonstrate: denoting with $\langle \cdot \rangle_{K_i}$ the average over the K_i's and with $\langle \cdot \rangle_N$ the average over N, we have

$$f_\Sigma(x) = \langle x^\Sigma \rangle = \left\langle \langle x^\Sigma \rangle_{K_i} \right\rangle_N = \left\langle f_K(x)^N \right\rangle_N = f_N(f_K(x)), \tag{5.14}$$

where the last equality is just the definition of the probability generating function of the random variable N, evaluated at $f_K(x)$. We stress that this is only valid for independent random variables.

5.2.3 Distribution of Number of Elements per Generation

Going back to the Galton-Watson branching process, where we know that $Z_{t+1} = \sum_{i=1}^{Z_t} K_i$, we can identify Z_{t+1} as Σ and Z_t as N; then equation (5.13) reads,

$$f_{Z_{t+1}}(x) = f_{Z_t}(f_K(x)) = f_{Z_t}(f(x)) \tag{5.15}$$

(dropping the subindex K). As $f_{Z_1}(x) = f(x)$, it is straightforward to see by induction that the generating function of Z_t, is given by

$$f_{Z_t}(x) = f(f(...f(x))) = f^t(x), \tag{5.16}$$

where the superindex t denotes composition t times. This is valid for $t = 1, 2, \ldots$; for $t = 0$ we have, obviously, that $f_{Z_0}(x) = x$ (because $Z_0 = 1$ with probability 1). In words, the generating function of the number of elements for each generation is obtained by the successive compositions of $f(x)$. This non-trivial result was first proved by Watson in 1874 (Harris 1963).

5.2.4 Expected Number of Elements per Generation

Here we present an illuminating result, which will be useful at some point in the chapter. Although, in general, the successive compositions of the generation function leads to very complicated mathematical expressions, the moments of Z_t can be computed in a simple way (Harris 1963). Using what we have learnt about generating functions together wtih equation (5.16), the expected value of Z_t is

$$\langle Z_t \rangle = \left. \frac{d}{dx} f^t(x) \right|_{x=1}. \tag{5.17}$$

Let us then write

$$\frac{d}{dx} f^t(x) = \frac{d}{dx} f(f^{t-1}(x)) = f'(f^{t-1}(x)) \frac{d}{dx} f^{t-1}(x), \tag{5.18}$$

therefore, by induction,

$$\frac{d}{dx} f^t(x) = f'(f^{t-1}(x)) f'(f^{t-2}(x)) \cdots f'(f^2(x)) f'(f(x)) f'(x). \tag{5.19}$$

Taking $x = 1$ and using that all the generating functions have to be 1 at that point,

$$\langle Z_t \rangle = f'(1)^t = m^t. \tag{5.20}$$

So, when $m < 1$ the mean number of elements per generation decreases exponentially, whereas when $m > 1$ this number increases, constituting a

stochastic realization of Malthusian growth. For this reason m is sometimes called the branching ratio. When $m = 1$ the average size of the population is constant, but we will later see that this does not mean that the population reaches a stable state. Higher-order moments can be computed in a similar way, but they are not so useful as the mean.

Another related issue is the one of the expected value of the number of elements per generation conditioned to the value of the previous generation, i.e., $\langle Z_{t+1}|Z_t = z_t\rangle$. As when Z_t is fixed, $Z_{t+1} = \sum_{i=1}^{z_t} K_i$, then, taking the expected value,

$$\langle Z_{t+1}|Z_t = z_t\rangle = \sum_{i=1}^{z_t}\langle K_i\rangle = z_t m. \tag{5.21}$$

This result can be used to relate branching processes with martingales (Grimmett and Stirzaker 2001), but this does not have to bother us.

5.2.5 The Probability of Extinction

Extinction of the process is achieved when $Z_t = 0$, for the first "time" (i.e., for the generation that yields $Z_t = 0$ for the first t). Then, all the subsequent Z's are also zero, and extinction can be considered an "absorbing state", in this sense. We now see that the probability of extinction in the Galton-Watson process is equal to one (extinction for sure) for $m \leq 1$ and is smaller than one for $m > 1$.

This result, which may be referred to as the Galton-Watson-Haldane-Steffensen (criticality) theorem, was first proved by J. F. Steffensen, in the 1930's (being unaware of the work by Galton and Watson, and later progress by Haldane). As Kendall (1966) pointed out, after then, the same theorem "was to be re-discovered over and over again, especially during the [Second World] War period, and no doubt we have not yet seen its last re-discovery". Ironically, Kendall did not know that Irénée-Jules Bienaymé knew the theorem, in its correct formulation, 30 years in advance Galton and Watson and 85 years before Steffensen (Kendall 1975)!

Indeed, extinction may happen at the first generation, $Z_1 = 0$, or at the second, $Z_2 = 0$, etc. All these extinction events are included in $Z_t = 0$, with $t \to \infty$; therefore, the probability of extinction P_{ext} is given by

$$\begin{aligned} P_{\text{ext}} &= \lim_{t\to\infty} P(Z_1 = 0 \text{ or } Z_2 = 0 \text{ or } \ldots \text{ or } Z_t = 0) = \\ &= \lim_{t\to\infty} P(Z_t = 0) = \lim_{t\to\infty} f^t(0), \end{aligned} \tag{5.22}$$

i.e., by the infinite iteration of the point $x = 0$ through the generating function $f(x)$ (using the key property that the probability of a zero value is the value of the generating function at zero, and equation (5.16) again).

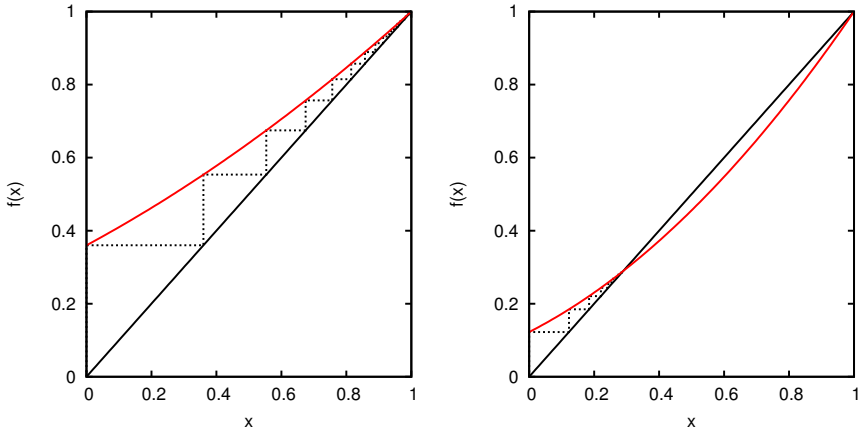


Fig. 5.4: Probability generating function $f(x)$ of the number of offsprings per element and iteration of the point $x = 0$ through successive compositions of f. The fixed points correspond to the crossings of the diagonal; the one closer to zero in each case is also the attractor for the iteration. Left corresponds to a subcritical case and right to a supercritical case. The model is the binomial one, with $n = 2$.

We now calculate the iteration $f^t(0)$. In the interval $[0, 1]$ the function $f(x)$ is non-decreasing and non-convex, taking values from p_0 to 1. If the slope of $f(x)$ at $x = 1$, given by $m = \langle K \rangle = f'(1)$, is smaller than or equal to 1, then $f(x)$ only crosses (or reaches) the diagonal at $x = 1$ (otherwise, $f(x)$ would need to be convex somewhere), and the iteration of the point $x = 0$ ends at the point $x = 1$ (which is the attractor, see Fig. 5.4). Therefore,

$$P_{\rm ext} = \lim_{t\to\infty} f^t(0) = 1, \tag{5.23}$$

i.e., extinction is unavoidable if $m \leq 1$. There is a trivial exception, though, associated to $p_1 = 1$ (and zero for the rest); this is an extremely boring situation indeed. In this case, $f(x) = x$, and therefore $\lim f^t(0) = 0$, which means, obviously, that the probability of extinction is zero.

If the slope of $f(x)$ at $x = 1$ is $m > 1$ (which only can happen for a non-linear generating function, $p_0 + p_1 < 1$), then $f(x)$ has to cross the diagonal at a point x^* smaller than one, which is the attractive solution to which the iteration tends, see Fig. 5.4 again. In mathematical language,

$$P_{\rm ext} = \lim_{t\to\infty} f^t(0) = x^*, \tag{5.24}$$

where

$$x^* = f(x^*) \text{ with } x^* < 1. \tag{5.25}$$

The demonstration is elaborated in the Appendix.

Summarizing,

$$P_{\text{ext}} = \begin{cases} 1 & \text{if } m \leq 1 \\ x^* & \text{if } m > 1 \end{cases} \tag{5.26}$$

with $x^* < 1$, except in the trivial case $p_1 = 1$, which has $m = 1$ but yields $P_{\text{ext}} = 0$.

Equation (5.26) clearly shows that, in general, the point $m = 1$ separates two distinct behaviors: extinction for sure for $m \leq 1$ and the possibility of non-extinction (non-sure extinction) for $m > 1$. Therefore, $m = 1$ constitutes a critical case separating these behaviors, called therefore subcritical ($m < 1$) and supercritical ($m > 1$). It is instructive to point out that, as $x = 1$ is always a solution of $f(x) = x$, Watson concluded, incorrectly, that the population always gets extinct, no matter the value of m (Kendall 1966).

5.2.6 The Probability of Extinction for the Binomial Distribution

For the sake of illustration we will consider a simple concrete example, a binomial distribution (Ross 2002, Grimmett and Stirzaker 2001),

$$p_k = P(K = k) = \binom{n}{k} p^k (1-p)^{n-k}, \text{ for } k = 0, \ldots n. \tag{5.27}$$

This assumes that each element has only a fixed number of trials n to generate other elements, and any of these n trials has a constant probability p of being successful. The generating function turns out to be, using the binomial theorem

$$f(x) = \sum_{k=0}^{\infty} \binom{n}{k} (1-p)^{n-k} p^k x^k = (1 - p + px)^n. \tag{5.28}$$

Let us consider the simple case with $n = 2$, and define $q = 1 - p$. As we know, the probability of extinction will come from the smallest solution in $[0, 1]$ of

$$x = (q + px)^2. \tag{5.29}$$

So,

$$x = \frac{1 - 2pq \pm \sqrt{(1-2pq)^2 - 4p^2q^2}}{2p^2}, \tag{5.30}$$

but the square root can be written as $\sqrt{1 - 4p(1-p)} = \sqrt{(1-2p)^2} = (1 - 2p)$, and then,

$$x = \frac{1 - 2p + 2p^2 \pm (1-2p)}{2p^2} = \begin{cases} \left(\frac{q}{p}\right)^2 \\ 1 \end{cases} \tag{5.31}$$

Therefore, the smallest root depends on whether p is below or above $1/2$

$$P_{\text{ext}} = \begin{cases} 1 & \text{for } p \leq \frac{1}{2} \\ \left(\frac{q}{p}\right)^2 & \text{for } p \geq \frac{1}{2} \end{cases} \tag{5.32}$$

As for the binomial distribution $m = np = 2p$ (Ross 2002), the critical case $m = 1$ corresponds obviously to $p = 1/2$, in agreement with the behavior of P_{ext}.

5.2.7 No Stability of the Population

Although this subsection contains an interesting result to better understand the behavior of the Galton-Watson process, it can be skipped as it is not connected to the rest of the chapter. In fact, the iteration of the point $x = 0$ shows what happens to the whole generating function of Z_t when $t \to \infty$. Indeed, in the same way as in subsection 2.5,

$$\lim_{t\to\infty} f_{Z_t}(x) = \lim_{t\to\infty} f^t(x) = 1 \text{ if } m \leq 1, \tag{5.33}$$

whereas

$$\lim_{t\to\infty} f_{Z_t}(x) = \lim_{t\to\infty} f^t(x) = x^* < 1 \text{ if } m > 1, \tag{5.34}$$

except for $x = 1$, which always fulfills $\lim_{t\to\infty} f^t(x) = 1$, see Fig. 5.5).

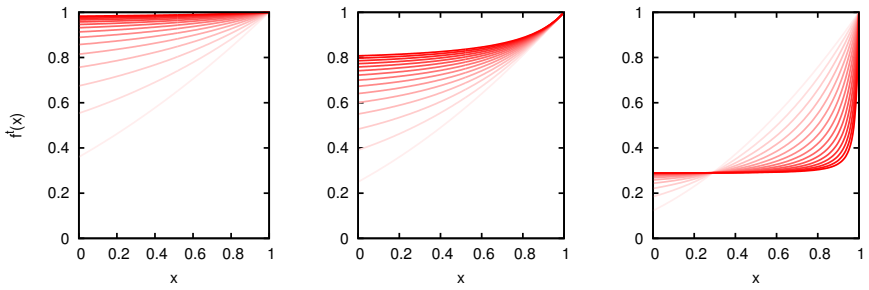


Fig. 5.5: Successive compositions of $f(x)$, for all x, yielding the probability generating functions of Z_t, starting at $t = 1$ (lighter red) up to $t = 15$ (darker red). Larger t leads to flatter functions, approaching the fixed point. From left to right, subcritical, critical, and supercritical cases, using a binomial model with $n = 2$.

Note that a flat generating function corresponds to probabilities equal to zero, except for the zero value, i.e.,

$$\lim_{t\to\infty} P(Z_t = k) = 0, \text{ except for } k = 0. \tag{5.35}$$

In this way, for $m \leq 1$ we have that $\lim_{t\to\infty} P(Z_t = 0) = 1$, and the population gets extinct; but for $m > 1$ we have found $\lim_{t\to\infty} P(Z_t = 0) = x^* < 1$; having any other finite value of K a zero probability, this means that Z_t goes to infinite, when $t \to \infty$, with probability $1 - x^*$; that is, Z_t cannot remain positive and bounded. The only stable state is extinction. Obviously, in this limit the Galton-Watson process is unrealistic, as other external factors should prevent that the population goes to infinity. But we do not need to bother about that, if we understand the limitations of the model.

5.2.8 Non-Equilibrium Phase Transition

Let us analyze in more detail what happens around the "transition point" $m = 1$. As we just have seen, recall equation (5.25), the extinction probability is given by the solution of $P_{\text{ext}} = f(P_{\text{ext}})$. When $m \leq 1$ the only solution in $[0, 1]$ is $P_{\text{ext}} = 1$ (except in the trivial case $p_1 = 1$). When $m > 1$ we have to take the smallest solution of $P_{\text{ext}} = f(P_{\text{ext}})$ in $[0, 1]$. In terms of the non-extinction probability, $\rho = 1 - P_{\text{ext}}$, we need to look for the largest ρ that is solution of

$$f(1-\rho) = \sum_{k=0}^{\infty} p_k (1-\rho)^k = 1 - \rho, \tag{5.36}$$

in the range $[0, 1]$. We explore the case of P_{ext} close to 1, for which ρ is close to zero, and, using the binomial theorem, we can expand $(1-\rho)^k = 1 - k\rho + k(k-1)\rho^2/2 + \cdots$, which yields

$$\begin{aligned} \sum_{k=0}^{\infty} p_k - \sum_{k=0}^{\infty} k p_k \rho + \frac{1}{2} \sum_{k=0}^{\infty} k(k-1) p_k \rho^2 + \cdots = \\ = 1 - m\rho + \frac{1}{2}\mu\rho^2 + \cdots = 1 - \rho, \end{aligned} \tag{5.37}$$

where we have introduced the mean m and the second factorial moment $\mu = \langle K(K-1) \rangle$ (which we assume exists). Therefore, up to second order in ρ we need to solve

$$\left(\frac{1}{2}\mu\rho + 1 - m \right) \rho \simeq 0. \tag{5.38}$$

It is immediate that one solution of equation (5.38) is $\rho = 0$, and one can realize that this solution is exact up to any order in ρ. The other solution is $\rho \simeq 2(m-1)/\mu$, but we must pay attention to the value of μ, which can

be written as $\mu = \sigma^2 + m(m-1)$, with $\sigma^2 = \langle (K-m)^2 \rangle = \langle K^2 \rangle - m^2$, i.e., the variance. Existence of m and σ^2 guarantees the existence of μ, then. Assuming $\sigma^2 \neq 0$,

$$\frac{2(m-1)}{\mu} = \frac{2(m-1)}{\sigma^2[1+m(m-1)/\sigma^2]} = \frac{2(m-1)}{\sigma^2}\left[1 - \frac{m(m-1)}{\sigma^2} + \dots\right] \tag{5.39}$$

(using the formula for the geometric series), therefore, ρ around zero means m around one, and we can write the second solution as

$$\rho \simeq \frac{2(m-1)}{\sigma^2} \tag{5.40}$$

which is only in the range of interest for $m > 1$.

In conclusion, we have

$$\begin{array}{ll} \rho = 0 & \text{if } m \leq 1 \\ \rho \simeq 2(m-1)/\sigma^2 & \text{if } m > 1, \end{array} \tag{5.41}$$

valid in the limit of small ρ. For $m > 1$ this limit is equivalent to $m \to 1$. The separate case $\sigma^2 = 0$ is only achieved in the trivial situation where $p_1 = 1$ (otherwise, the mean cannot approach one).

In this way, we obtain a behavior that is the one corresponding to a continuous phase transition in thermodynamic equilibrium. Identifying m with a control parameter (as temperature, or more properly, the inverse of temperature) and ρ with an order parameter (as magnetization in a magnetic system) these transitions show an abrupt but continuous change of ρ as a function of m at the transition point m_c, with

$$\begin{array}{ll} \rho = 0 & \text{below } m_c \\ \rho \propto (m - m_c)^\beta & \text{above but close to } m_c \end{array} \tag{5.42}$$

For magnetic systems, m_c corresponds to the so-called Curie temperature. For the Galton-Watson branching process we can extract from equation (5.41) that

$$m_c = 1 \text{ and } \beta = 1, \tag{5.43}$$

where we assume that the variance of K does not go to zero at the transition point.

We can compare the previous general result, $\rho \simeq 2(m-1)/\sigma^2$, for m above but close to 1, with the result we found for the binomial distribution with $n = 2$ (see equation (5.32)), for which

$$\rho = 1 - \left(\frac{1-p}{p}\right)^2 = \frac{2p-1}{p^2} \tag{5.44}$$

when $p \geq 1/2$. Using that in this case $m = np$ and $\sigma^2 = npq$ (see Ross (2002)),

$$\frac{2(m-1)}{\sigma^2} = \frac{2p-1}{pq} \simeq \frac{2p-1}{p^2}, \tag{5.45}$$

because $q = 1 - p \simeq p$ for $p \simeq 1/2$. So, equations (5.32) and (5.41) agree close to the transition point. Figure 5.6 shows also how they disagree as m increases.

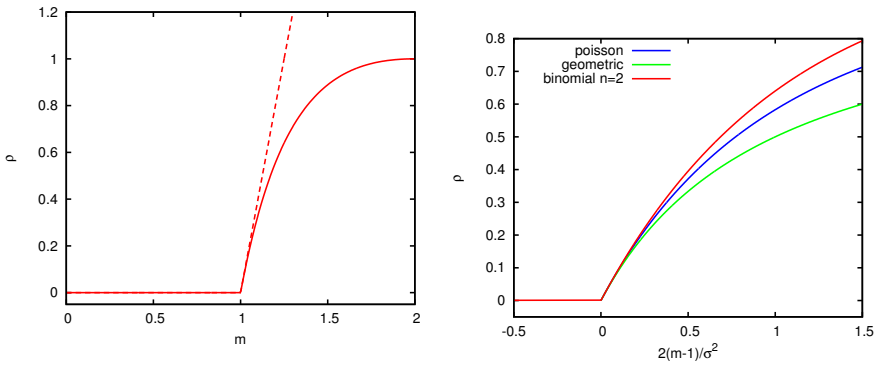


Fig. 5.6: Left: non-extinction probability ρ as a function of the mean number of offsprings per element, m. Dashed line corresponds to the approximation explained in the text (eq. (5.41)). The abrupt change in ρ is the hallmark of a continuous phase transition. The model is binomial with $n = 2$. Right: the same but as a function of the rescaled distance to the critical point, $2(m-1)/\sigma^2$, where σ^2 refers to the variance at $m = 1$. The Poisson and the geometric distributions are also studied.

Finally, for completeness, we can play with the pathological case given by $\sigma^2 = 0$. Let us consider first the following model, $p_0 = 1 - \lambda_1$, $p_1 = \lambda_1$ (and zero otherwise), with $\lambda_1 < 1$. Then, $m = \lambda_1$, and we know that $\rho = 0$. Next, let us consider $p_1 = 1 - \lambda_2$, $p_2 = \lambda_2$ (and zero otherwise), giving $m = 1 + \lambda_2$. In this case, $\rho = 1$ always, yielding a discontinuous, or first order phase transition.

5.2.9 Distribution of the Total Size of the Population: Binomial Distribution and Rooted Trees

Our main interest will now be to calculate the total size S of the population, summing across all generations, i.e.,

$$S = \sum_{t=0}^{\infty} Z_t, \tag{5.46}$$

this corresponds to the total number of individuals that have ever been born, the total number of neutrons participating in a nuclear chain reaction, or the energy released during an event in an earthquake model.

Let us go back to the concrete binomial case,

$$p_k = P(K = k) = \binom{n}{k} p^k (1-p)^{n-k}, \text{ for } k = 0, \ldots, n. \tag{5.47}$$

The size distribution can be calculated using elementary probability and combinatorics. One needs to take advantage of the representation of a branching process as a tree (which is a connected graph with no loops). Each element is associated to a node, and branches linking nodes indicate an offspring relationship between two nodes. Naturally, all nodes have just one incoming branch, except the one corresponding to the zero generation (which in this context is called the root of the tree). So, the number of branches is the number of nodes minus 1. As the size s of a tree is the number of nodes it contains, the number of branches is $s-1$, and the number of missing branches (non-successful reproductive trials) is $ns-(s-1)$ (because the number of possible branches arising from s nodes is ns) (Christensen and Moloney 2005). Therefore, a particular tree of size s comes with a probability $p^{s-1}(1-p)^{(n-1)s+1}$, and the probability $P(S=s)$ of having an undefined tree of size s is obtained by summing for all possible trees of size s. In the case $n=2$ the number of trees with s nodes is given by the Catalan number

$$C_s = \frac{1}{s+1}\binom{2s}{s}, \tag{5.48}$$

see the Appendix for its calculation. Then,

$$P(S=s) = \frac{1}{s+1}\binom{2s}{s} p^{s-1}(1-p)^{s+1} \text{ with } s = 1, 2, \ldots \tag{5.49}$$

It can be checked, using the generating function of the Catalan numbers, that this expression is normalized for $p \leq 1/2$ but not for $p > 1/2$, in fact,

$$\sum_{s=1}^{\infty} P(S=s) = P_{\text{ext}}, \tag{5.50}$$

see the Appendix again.

Nevertheless, the exact expression we have obtained for $P(S=s)$ does not teach us anything about the behavior of this function (unless one has a great intuition about the behavior of the binomial coefficients). In this regard, Stirling's approximation is of great help (Christensen and Moloney 2005). It

states that, in the limit of large N one can make the substitution

$$N! \sim \sqrt{2\pi N}\left(\frac{N}{e}\right)^N, \tag{5.51}$$

see the Appendix once more. The symbol e is nothing else than the e number. So, for large sizes we can apply the approximation to s and also to $2s$,

$$(2s)! \sim \sqrt{4\pi s}\left(\frac{2s}{e}\right)^{2s}. \tag{5.52}$$

Therefore, the binomial coefficient turns out to be,

$$\binom{2s}{s} = \frac{(2s)!}{s!s!} \sim \frac{1}{\sqrt{\pi s}}\frac{(2s)^{2s}}{s^{2s}} \sim \frac{4^s}{\sqrt{\pi s}}, \tag{5.53}$$

and the Catalan number, replacing $s+1 \sim s$,

$$C_s = \frac{1}{s+1}\binom{2s}{s} \sim \frac{4^s}{\sqrt{\pi}s^{3/2}}. \tag{5.54}$$

This is an exponential increasing function of s, and the term $s^{3/2}$ does not seem to play any role, asymptotically. However, introducing the factor $p^{s-1}(1-p)^{s+1}$, we go back to equation (5.49), getting

$$P(S=s) \sim \frac{1-p}{\sqrt{\pi}p}\frac{[4p(1-p)]^s}{s^{3/2}}. \tag{5.55}$$

Notice that $p(1-p)$ is no larger than $1/4$, so the exponential term becomes decreasing, except for $p = 1/2$, where it disappears. We can go one step further, by writing,

$$[4p(1-p)]^s = e^{s\ln[4p(1-p)]} = e^{-s/\xi(p)} \tag{5.56}$$

with the characteristic size defined as

$$\xi(p) = \left(\ln\frac{1}{4p(1-p)}\right)^{-1}, \tag{5.57}$$

and finally equation (5.55) reads,

$$P(S=s) \sim \frac{1-p}{\sqrt{\pi}p}\frac{e^{-s/\xi(p)}}{s^{3/2}}, \tag{5.58}$$

So, for s large, but substantially smaller than $\xi(p)$, the size probability mass function is a power law, with exponent $3/2$. For larger s, the exponential decay dominates. The exception is the critical case, $p = 1/2$, for which $\xi(p)$

becomes infinite, the exponential disappears and the distribution is a pure power law. In this case the exponent 3/2 is a critical exponent. The reader can see the goodness of the approximation in Fig. 5.7.

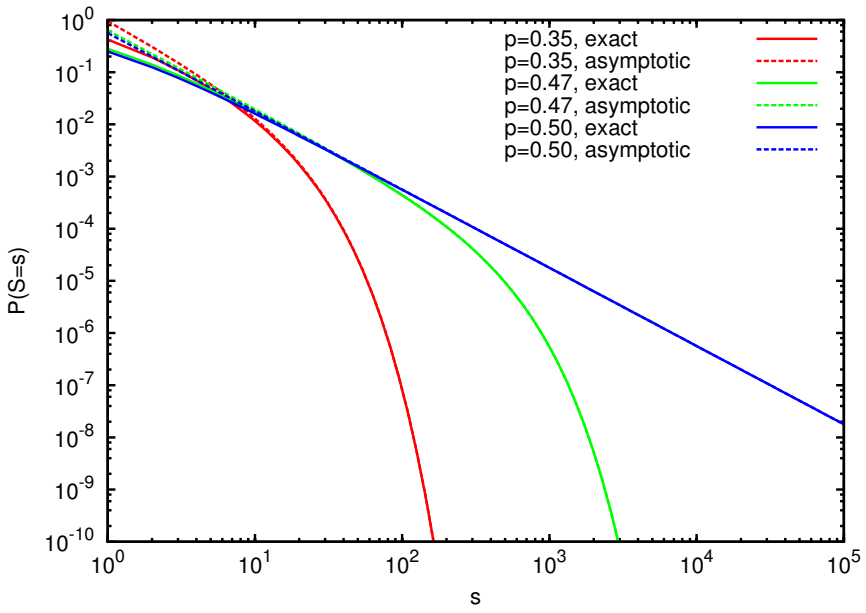


Fig. 5.7: Probability mass functions of the total size of the population S , for different values of the parameter p of a binomial distribution with $n = 2$, both in the subcritical and critical cases. The asymptotic solution for large s is also shown. The pure power law at the critical point becomes apparent.

Another critical exponent arises for the divergence of the characteristic size $\xi(p)$. Introducing the deviation with respect to the critical point, $\Delta \equiv p - p_c = p - 1/2$, one can write,

$$p(1-p) = \frac{1}{4} - \Delta^2, \tag{5.59}$$

and so, close to the critical point (for small Δ),

$$\frac{1}{4p(1-p)} = \frac{1}{1-4\Delta^2} \simeq 1 + 4\Delta^2 + \ldots \tag{5.60}$$

(using the formula of the geometric series), then

$$\ln \frac{1}{4p(1-p)} \simeq \ln(1+4\Delta^2) \simeq 4\Delta^2 + \ldots \tag{5.61}$$

(using the Taylor expansion of the logarithm at point 1) and

$$\xi(p) = \left(\ln \frac{1}{4p(1-p)} \right)^{-1} \simeq \frac{1}{4\Delta^2} + \dots \tag{5.62}$$

Therefore, the characteristic size $\xi(p)$ diverges at the critical point as a power law, with an exponent equal to 2. This allows to write the asymptotic formula (s large) for the size distribution in a simpler form, close to the critical point (Δ small),

$$P(S = s) \sim \frac{1-p}{\sqrt{\pi}p} \frac{e^{-4(p-p_c)^2 s}}{s^{3/2}}. \tag{5.63}$$

Hence, after this perhaps long but worthwhile digression, we are able to say something about the energy distribution in Otsuka's model, which the reader will have already noted is a particular case of the Galton-Watson process. If one takes $p < 1/2$ the resulting energy distribution has an exponential tail, with a characteristic scale given by $\xi(p)$. This means that earthquakes attenuate, or get extinct, and in no way can dissipate energies larger than the scale provided by $\xi(p)$ (the probability of having an earthquake of size larger than $10\xi(p)$ is ridiculously small). This is the subcritical case. On the other hand, if $p > 1/2$ there are two types of earthquakes, first, those similar to the subcritical ones, with a size limited by the scale defined by $\xi(p)$, and second, infinite or never-ending earthquakes ($P_{\text{ext}} < 1$), where the initial small perturbation (the toppling of just one domino piece) grows exponentially. This is the supercritical regime (Ben-Zion 2008). Neither the subcritical nor the supercritical case are in correspondence with the Gutenberg-Richter law, which yields a power-law distribution of energies, and therefore the absence of a characteristic scale. But this is precisely what corresponds to the critical case, $p = 1/2$, which yields also a power-law distribution. Thus, the propagation of an earthquake through a fault is not only stochastic in the sense that when a patch fails one does not know what will happen next, but it is worse than that, as a critical process is equally likely to intensify or attenuate. Note how difficult is to achieve a critical behavior, as p has to be finely tuned to $1/2$, otherwise criticality is lost. This is also what is really difficult in terms of domino topplings, and not to get a full-system supercritical toppling that could break the Guinness world record (which is trivial from a mathematical point of view).

The agreement between the model and real earthquakes is qualitative but not quantitative, as the model leads to $\alpha = 3/2$ whereas for earthquakes $\alpha \simeq 5/3 \simeq 1.67$. In the next subsection we will explain that the model value of $3/2$ is rather robust and other versions of the Galton-Watson process lead to the same exponent. This discrepancy has been explored in detail by Kagan (2010), who argues that there are a series of technical artifacts that make increase the value of the exponent for earthquakes, and therefore, following Kagan, both exponents would be close and probably compatible.

5.2.10 Generating Function of the Total Size of the Population

In order to advance further in the understanding of branching processes, our little story carries us to the U.S. during the Second World War. While soldiers were fighting in the field and civilians were suffering the horrors of war, a group of scientists gathered in the peace of Los Alamos, New Mexico, to do research to develop the first nuclear bombs. Among these brilliant people was the great Polish mathematician Stanislaw Ulam, who was hired by his famous colleague John Von Neumann (Ulam 1991). Together with David Hawkins (philosopher of science and most talented amateur mathematician ever known by Ulam) they were investigating the multiplication of neutrons in nuclear chain reactions, using what we call now branching processes. It seems that they were unaware of the pioneering work of Galton and Watson.

Hawkins and Ulam showed, among other things, that the generating function $g(x)$ of the total size of the population, $S = \sum_{\forall t} Z_t$, fulfills, in the non-supercritical case,

$$g(x) = xf(g(x)) \tag{5.64}$$

where, as usual, $f(x)$ is the generating function of the number of offsprings of an individual element. What follows in this subsection is based in their work for the Manhattan Project (Hawkins and Ulam 1944, Ulam 1990), but our derivation is somewhat simpler. What we call total size of the population will correspond to all neutrons generated during the reaction.

First, it is convenient to consider the size from generation 1 to τ (excluding by now the zero generation). This is

$$S_\tau = \sum_{t=1}^{\tau} Z_t \tag{5.65}$$

with probabilities $q_s^{(\tau)} = P(S_\tau = s)$ and a generating function $\tilde{g}_\tau(x) = \sum_{\forall s} q_s^{(\tau)} x^s$. A size s in generations from 1 to τ can be decomposed into a size k in the first generation, with probability p_k, and a size $s-k$ in the remaining $\tau - 1$ generations (from 2 to τ), but starting with k elements; this has a probability $q_{s-k}^{(\tau-1,k)}$. (Note that, with this notation $q_s^{(\tau)} = q_s^{(\tau,1)}$.) Then, using the law of total probability,

$$q_s^{(\tau)} = \sum_{k=1}^{s} p_k q_{s-k}^{(\tau-1,k)}, \tag{5.66}$$

except for $s = 0$, where $q_0^{(\tau)} = p_0$. If we multiply by x^s and sum for all s, from 0 to ∞, we will obtain on the left hand side the generating function of S_τ, which turns out to be

$$\tilde{g}_\tau(x) = p_0 + \sum_{s=1}^{\infty} \sum_{k=1}^{s} p_k q_{s-k}^{(\tau-1,k)} x^s = p_0 + \sum_{k=1}^{\infty} p_k \left[\sum_{s=k}^{\infty} q_{s-k}^{(\tau-1,k)} x^{s-k} \right] x^k. \quad (5.67)$$

The term inside the square brackets is the generating function of the size from 1 to $\tau - 1$ generations but, instead of starting with one single element (the usual $Z_0 = 1$), starting with k elements ($Z_1 = k$). As these k parents are independent of each other, the resulting size will be the sum of k independent random variables, each with generating function $\tilde{g}_{\tau-1}(x)$, which yields $[\tilde{g}_{\tau-1}(x)]^k$ as the corresponding generating function, that is,

$$[\tilde{g}_{\tau-1}(x)]^k = \sum_{s-k=0}^{\infty} q_{s-k}^{(\tau-1,k)} x^{s-k}, \quad (5.68)$$

Substituting into equation (5.67), this leads to

$$\tilde{g}_\tau(x) = p_0 + \sum_{k=1}^{\infty} p_k [\tilde{g}_{\tau-1}(x)]^k x^k = f(x\tilde{g}_{\tau-1}(x)) \quad (5.69)$$

where we have introduced the definition of $f(x) = f_K(x)$.

If we want to include the zero generation in the size, we need to add an independent variable with generating function x (as Z_0 takes the value 1 with probability 1), and then, the generating function of the size from generation 0 to τ is the product $g_\tau(x) = x\tilde{g}_\tau(x)$. This leads to

$$g_\tau(x) = xf(g_{\tau-1}(x)). \quad (5.70)$$

Coming back to the total size,

$$S = \sum_{t=0}^{\infty} Z_t, \quad (5.71)$$

the corresponding generating function is $g(x) = \lim_{\tau\to\infty} g_\tau(x)$. If the probability of extinction is one, i.e., if the system is not supercritical, this is the same as $\lim_{\tau\to\infty} g_{\tau-1}(x)$, and therefore we have

$$g(x) = xf(g(x)). \quad (5.72)$$

So, the desired generating function is the solution of this equation, with $f(x)$ known. We will not be able to solve it in general; however, notice that this is not necessary in order to get the moments of S. Differentiating equation (5.72) with respect x one obtains

$$g'(x) = f(g(x)) + xf'(g(x))g'(x), \quad (5.73)$$

and taking $x = 1$ and isolating,

$$\langle S\rangle = g'(1) = \frac{1}{1-f'(1)} = \frac{1}{1-m}, \tag{5.74}$$

which goes to infinity as $\langle K\rangle = m = f'(1)$ goes to 1, that is, at the critical point. Of course, as we have mentioned, the result is not applicable in the supercritical case, $m > 1$, where the population can growth to infinite with a non-zero probability. Further differentiation yields higher-order moments.

The same result could have been obtained directly, as

$$\langle S\rangle = \langle Z_0+Z_1+Z_2+\cdots\rangle = \langle Z_0\rangle+\langle Z_1\rangle+\langle Z_2\rangle+\cdots = 1+m+m^2+\cdots = \frac{1}{1-m}, \tag{5.75}$$

where the last equality only holds in the subcritical case, otherwise, $\langle S\rangle$ goes to infinity.

In a few cases, the equation for $g(x)$ allows to easily obtain a solution. Revisiting the binomial example with $n = 2$, for which $f(x) = (1-p+px)^2$, one gets

$$g(x) = xf(g(x)) = x(1-p+pg(x))^2, \tag{5.76}$$

from where

$$g(x) = \frac{1-2pqx \pm \sqrt{1-4pqx}}{2p^2x}, \tag{5.77}$$

with $q = 1-p$. Using the Taylor expansion for the square root term (see the Appendix),

$$\sqrt{1-4pqx} = 1-2pqx-\sum_{s=1}^{\infty}\frac{(2s-1)!!2^{s+1}}{(s+1)!}(pqx)^{s+1}, \tag{5.78}$$

and recognizing the Catalan numbers C_s there, we get (see the Appendix),

$$g(x) = \frac{q}{p}\sum_{s=1}^{\infty} C_s(pqx)^s, \tag{5.79}$$

where we also realize that only the minus sign before the square root leads to a true generating function. Therefore, the coefficients of x^s lead to

$$P(S=s) = C_s p^{s-1}q^{s+1}, \tag{5.80}$$

for $s \geq 1$. This result is exactly the same as the one we obtained previously in a different manner (see equation (5.49)), although in this way we do not need to count trees, as the Catalan numbers arise directly in the series expansion (in fact, we do not even need to know them).

We confirm that the results for Otsuka's binomial model yield a size exponent equal to 3/2. But it would be desirable to test the robustness of such exponent value, as, after all, the model is a crude simplification of reality, and we would like that modifications of the model do not lead to a totally

different behavior. Despite the difficulty to find the power-law behavior (for which we need to finely tune the parameter p to $1/2$), if one considers other models different than the binomial one, the asymptotic behavior of the size distribution is in general always given by a power law with exponent $3/2$, in the critical case; this can be proved by means of Cauchy's formula and assuming only finite variance, see Otter (1949), Harris (1963). So, going beyond robustness, it is common to denote such invariance as universality.

5.2.11 Self-Organized Branching Process

At this point we are ready to accept the agreement, not only qualitative but, following Kagan's remarks (Kagan 2010), also quantitative, between a critical branching process and earthquake occurrence. So, in order to tune the model to reality we just need to take $p = 1/2$ (in Otsuka's binomial case) or $m = 1$ (in general) and the agreement is really satisfactory, and we could finish our search for a model here.

But we can try to go one step farther and ask: why do we find that the tectonic systems (and other geosystems related to natural catastrophes) are always keeping a delicate balance between a subcritical and a supercritical state, i.e., in an apparent critical state? Can the coincidence be just fortuitous? In the reproduction of individuals one could devise an evolutionary explanation. Imagine a series of isolated islands, each one occupied by a population following a Galton-Watson process but with different parameters for each island. It is clear that islands with subcritical populations get deserted after a number of generations. Populations in supercritical islands either get extinct also or explode exponentially, in which case we assume that the population collapses, due to the exhaustion of the resources (this is an ingredient that is not in the original Galton-Watson model). In the critical case, the population also gets extinct, but for a few of these islands the population can survive for very long times, much longer than in the subcritical and supercritical cases. So, after a long enough time we would only find critical populations.

However, this evolutionary scenario is not applicable to a tectonic system, where, when the process (the earthquake) gets extinct, a new one will start sooner or later. Rather, the situation would be analogous to finding all magnetic materials on Earth at the onset of magnetization, which would mean that their temperatures would be equal to the Curie temperature of each material. One could suspect then that there is some mechanism enforcing criticality, where the temperature changes as a function of magnetization, and magnetization is kept at the border of the transition; in other words, both parameters are linked through some feedback mechanism (Sornette 1992, Pruessner and Peters 2006).

Zapperi et al. (1995) propose a model in this line. They start with a standard branching process but introduce some important modifications:

- They limit the number of generations to a maximum τ, so $0 \leq t \leq \tau$.
- After the extinction of the process (which is obviously certain when the number of generations is limited), the parameters of the process change for the next realization, in such a way that for subcritical cases ($m < 1$), the mean m of the number of offsprings for each individual unit increases, whereas in the supercritical case ($m > 1$) the mean m decreases. The idea is to make the critical state $m = 1$ an attractor of the dynamics.

In order to be more concrete, let us consider the usual binomial distribution with only 0, 1, or 2 possible offsprings and a probability p that each reproductive trial is successful. Then we already know that $p < 1/2$, $p = 1/2$, and $p > 1/2$ correspond to the subcritical, critical, and supercritical cases, respectively. The dynamics proposed by Zapperi and coauthors relies on the activity that reaches the "boundary" of the system (defined by the last generation, $t = \tau$), which is Z_τ, changing the probability p through the following formula

$$p(T+1) = p(T) + \frac{1 - Z_\tau(p(T), T)}{N}, \tag{5.81}$$

with T a discrete time index counting the number of realizations of the process (do not confuse with t) and $N = 2^{\tau+1} - 1$ the maximum number of possible elements, i.e., the number of branches of the underlying complete tree. Thus, if the activity does not reach the boundary, Z_τ is zero and the parameter p is increased by $1/N$, this is a very small number in the limit of very large systems ($N \to \infty$). On the other hand, if the activity at the boundary is greater than one, p is decreased by $(Z_\tau - 1)/N$.

We already know that the expected value of Z_τ is m^τ, with m the mean of the offspring distribution ($m = 2p$ in our particular binomial model). Let us introduce a noise term, η, which takes into account the fluctuations of Z_τ with respect its mean, i.e., $\eta = Z_\tau - m^\tau$. Obviously, by construction, $\langle \eta \rangle = 0$. If we neglect, for a while, the noise term in equation (5.81), the deterministic part reads,

$$p(T+1) = F(p(T)) = p(T) + \frac{1 - (2p(T))^\tau}{N}. \tag{5.82}$$

This is a discrete dynamical system, or a map, for which a fixed point $p^* = F(p^*)$ exists, $p^* = 1/2$. Moreover, the fixed point is attractive, as $|F'(p^*)| < 1$ (Alligood et al. 1997), due to $\tau \ll N$.

Taking into account the value of the standard deviation of Z_τ (Harris 1963), it can be shown that the noise term η/N will have a vanishing effect in the limit of very large systems, and then the stochastic evolution will lead the system towards the deterministic fixed point, plus small random fluctuations around it.

This spontaneous evolution of a system towards a particular organized state is referred to as self-organization. It is clear now that what Zapperi et al.

introduced is a branching process that self-organizes towards a critical state. Nevertheless, the particular dynamics they propose seems a bit arbitrary. How can this kind of global control be implemented in a real system, where we expect the interactions between elements to be purely local?

5.2.12 Self-Organized Criticality and Sandpile Models

In fact, the self-organized branching process introduced by Zapperi et al. (1995) was naturally embedded in the previous notion of self-organized criticality (SOC), invented by Bak and coworkers in the 1980's (Bak 1996, Jensen 1998, Christensen and Moloney 2005). Although it is not relevant for our story, it is worth to state that these authors were not interested in (because they were not aware of) the problem of power-law distributions in natural hazards (Bak 1996); rather, they were mainly concerned to similar-in-spirit problems in condensed-matter physics, as charge density waves and one-over-f noise, as well as to the emergence of fractal spatial structures elsewhere (Bak et al. 1987). The fact that earthquakes (and other hazards) were a manifestation of self-organized criticality was a fortunate by-product, pointed by Ito and Matsuzaki (1990), Sornette and Sornette (1989), and Bak and Tang (1989) shortly after the introduction of the SOC concept, see also the review of Main (1996). Nowadays, natural hazards are one of the main applications of SOC, despite the original lack of attention by Bak et al. (1987). As we have seen through this chapter, ignorance seems a common characteristic of science evolution.

The metaphor used by Bak in order to illustrate his ideas was that of a pile of sand (Bak 1996). We have to recognize that the sandpile we are going to consider is a bit esoteric; in fact, there is a clear correspondence between the model and a pile only in one dimension (the one-dimensional model corresponds to a pile constrained in two dimensions, between two parallel plates (Christensen et al. 1996)). But instead of keeping close to reality, it is more effective to deal with a mean-field sandpile; this is achieved either in a system defined in the limit of infinite dimensions or in a system in which each element has "random neighbors", and neglecting the correlations between the elements. Notice that Bak and colleagues make use of a new concept, not present in the branching processes already explained: the notion of complexity, understood here as the nontrivial interaction between many units or agents, which will result in an emergent collective behavior that is different than the sum of the behavior of the individual parts (Newman 2011).

So, consider a system consisting in a large number of elements, such that each element can store a certain number of discrete packages (or particles), but when this limit is surpassed the packages are released to other elements – the neighbors. The situation is analogous to what happens in a Ministry office. Each bureaucrat has a series of documents or papers (the packages) at his/her

desk, but when the number of those is too big, he/she decides to do something about it and transfers some papers to some other (random) bureaucrats, and so on (Bak 1996). This simple behavior will lead to interesting dynamics, unexpectedly.

To be specific, let us consider that each element can store at most one package; if some extra package arrives to it, the element releases two packages to some other units, taken randomly (either among all other elements, what defines random neighbors or among the $2d$ nearest neighbors in a $d-$dimensional square lattice). If, after the release, the number of packages is still greater than one (which may happen if the element received more than one package) the release process is repeated. All the elements evolve following a parallel updating of their dynamics, i.e., there is a common clock setting the time t of all elements. In a formula,

$$\text{if } z_i \geq 2 \Rightarrow \begin{cases} z_{n(i)} \to z_{n(i)} + 1, \\ z_i \quad\;\; \to z_i - 2, \end{cases} \tag{5.83}$$

where z_i counts the number of packages of element i and $n(i)$ denotes two of its neighbors.

Obviously, this process can give rise to an avalanche in the transference of packages, which only stops when all elements have no more than one package. In that case, the system is perturbed by the addition of one extra package to a randomly chosen element, and the dynamics starts again. This defines a new time scale, denoted by T (in the same way as in the previous subsection). So,

$$\text{if } z_i \leq 1, \forall i \Rightarrow\; z_j \to z_j + 1, \tag{5.84}$$

where j denotes a randomly selected unit. The system also releases packages outside (or to the garbage can, in the bureaucrats picture); in a $d-$ dimensional lattice this happens when a boundary element selects as a neighbor an external element; in a fully random-neighbor system this happen just with a small predefined probability for each element. This simple variation of the original sandpile model of Bak et al. (1987) (changing the topology of the system by means of a different selection of neighbors) can be viewed also as a mean-field version of the so-called Manna model (Manna 1991, Christensen and Moloney 2005).

The simple rules of the model make that the total number of packages in the system, M, evolves, from the addition of one package to the next, accordingly to

$$M(T+1) = M(T) + 1 - \text{drop}(T), \tag{5.85}$$

where drop is the number of packages that are expelled from the system. The key parameter of this model is p, defined, for each element, as the probability that its number of packages is equal to one (so they are at the onset of instability). But in a mean field description all elements are uncorrelated and equivalent, so we can define a generic p for the whole system, verifying

$p = M/N$, with N the total number of elements. So, there is a probability p that an element releases two packages when it receives one. The action of release is what constitutes the generation of an offspring, which is the element that relaxes. Therefore, dividing equation (5.85) by N we obtain

$$p(T+1) = p(T) + \frac{1 - \mathrm{drop}(T)}{N}, \tag{5.86}$$

which we can recognize as essentially equation (5.81), the one introduced by Zapperi et al. (1995) in the self-organized branching process. We have already realized that this equation provides a feedback mechanism of the number of packages into the toppling (branching) probability (early identifications of this obvious feedback in SOC were written by Kadanoff (1991) and Sornette (1992)).

Both in the limit of an infinite dimension lattice or in a fully random neighbor system one realizes that the evolution of an avalanche corresponds to a set of propagating non-interacting packages (as the probability that the activity comes back to an element is vanishingly small), and therefore the activity evolves as a branching process. But note that the tree associated to the branching process does not correspond to a quenched underlying structure of the system, as the random neighbors are selected dynamically, at each time step. The limit τ in the number of generations introduced by Zapperi and coauthors needs to be added as an extra ingredient in the model, enforcing the dissipation of packages to take place at the τ time step. In summary, this illustrates the correspondence between the mean-field limit of sandpile models and branching processes. This is enough for our purposes. Other chapters in this book illustrate in much more detail the dynamics of sandpiles. Nevertheless, it is worth mentioning that the first connection between SOC and critical branching process was published by Alstrøm (1988), where it was assumed, however, that the system was in a critical state from the beginning. Notably, much before, Vere-Jones (1976) had proposed a branching model very similar to Otsuka's (but, as usual, Vere-Jones was unaware of Otsuka's work) and realized that the tectonic system should evolve spontaneously towards criticality. Also, very recently, Hergarten (2012) has introduced a variation of Zapperi et al.'s branching model that evolves only with local rules.

Recapitulating, self-organized criticality offers a coherent framework for the understanding of earthquakes and many other natural hazards mentioned in the first section. Indeed, both phenomena (SOC and earthquakes) show a highly non-linear response, where a small and slow perturbation or driving (the addition of grains, or the stress provided by the motion of the tectonic plates) pumps energy into the system, which, due to the presence of local thresholds stores that energy, until at some point some threshold is surpassed. The resulting release of energy propagates locally, which can trigger further surpassings of thresholds, generating a chain reaction or avalanche. One key point is that the energy released in such a way has to be power-law distributed, so the system responds in all possible scales. Notice also

that the dynamics shows a time-scale separation, as the avalanches happen infinitely fast compared with the driving (the toppling of grains is stopped during the propagation of an avalanche). Moreover, Main (1996) mentions additional characteristics of seismicity present in SOC models, namely, stress drops that are small in comparison with the regional tectonic stress field and the existence of seismicity induced or triggered by relatively small stress perturbations. All this makes SOC a very plausible mechanism for earthquakes. The connection is made still more concrete using variations of the sandpile models that mimic the behavior of the spring-block model of Burridge and Knopoff (1967) as the so-called OFC model (Olami et al. 1992). See also Main (1996).

However, as far as we know, the authentic hallmark of SOC, the existence of an underlying second-order (continuous) phase transition, has not been found in earthquakes. The very nature of SOC makes almost impossible to identify such an abrupt change of an order parameter when a control parameter changes (because the control parameter is attracted towards the critical point). Nevertheless, this elusive behavior has been found in a different system: rainfall (Peters and Neelin 2006), thanks to very large fluctuations from criticality; so, if a control and an order parameter could be measured and if similarly large fluctuations were exist, one would finally prove the existence of SOC in earthquakes.

The same reasoning applies to other natural hazards, for which, at least, sandpile-like models are abundant in the literature, and their classification as SOC systems is plausible (Jensen 1998). The case of hurricanes is still not clear (Corral 2010), whereas for tsunamis we can state that their power-law distribution (Burroughs and Tebbens 2005) does not arise from a SOC mechanism, as they are not slowly driven (rather, they are violently driven by earthquakes, landslides and meteorite impacts).

Finally, it is worth mentioning that there is another connection between branching processes and earthquakes. Instead of using the branching to model the propagation of individual earthquakes, it is used for the way in which one earthquake triggers other earthquakes, i.e., aftershocks, following the so-called Omori law. The most representative model of this kind is the epidemic-type aftershock-sequences (ETAS) model (Ogata 1999, Helmstetter and Sornette 2002). Interestingly, the evolution model of Bak and Sneppen (1993) (another paradigm of SOC) can be interpreted to reproduce the statistics of earthquakes from this (slow) time scale (Ito 1995). This perspective opened a whole new line in statistical seismology, but this is a different story (Bak et al. 2002, Corral 2004a,b).

5.3 Conclusions

We started this chapter showing some remarkable statistical properties of earthquake occurrence, and ended up mingling with infinite-dimensional sandpiles models for self-organized criticality. In between, we learnt a few things about branching processes. Now we sketch some consequences for our initial object of study: natural hazards.

First, besides any model, we can say a few things just by looking at the data: earthquakes and other natural hazards follow a power-law distribution of sizes, in some cases with an exponential cutoff due to finite-size effects (the Earth is finite, after all!). For the particular values of the exponents found, this implies that, although big events are less likely, they are always the main contributors of the overall devastation. As financial data of asset returns and other social and technological data have also been reported to follow power law distributions (Mantegna and Stanley 1999, Newman 2005), one wonders what the points in common with these systems and natural hazards can be.

Regarding Otsuka's rupture model, we showed how, by using a fairly simple stochastic cascade setup for the local dynamics of fault patches and the mathematical formalism for branching processes, one can reproduce the global statistical properties of real earthquake occurrences (and other natural hazards). This is quite remarkable, as it constitutes a link between two distinct observational scales: the micro-scale of local dynamics, and the macro-scale of global statistical behavior.

But Otsuka's model is a particular case of the Galton-Watson branching process. So, first, we presented in an easy way the main results already known for such processes (main results in relation to our interests). We explained how the machinery of probability generating functions allows to find a formula for the activity (or population) at any generation of the process. In the limit of infinite generations, one gets the probability of extinction, which shows an abrupt change between two different regimes: extinction for sure if the mean number of offsprings is below or equal to one, and the possibility of non-extinction in the opposite case. Further progress leads to an expression for the probability of the total size of the process (the total population ever born or the total energy radiated by an earthquake). It is precisely at the border of the two mentioned cases, at the critical point of the transition, that one finds a behavior compatible with earthquakes and other natural hazards. A power-law distribution with exponent 3/2 emerges in this case; however, it remained unexplained how the Earth should drive itself towards such a critical state.

In this regard, we showed how, by using a simple feedback mechanism, one can turn the critical point into an attractor of the model. A global condition, related with boundary dissipation, acts on the probability of activation, in such a way that when this probability is low, it increases, and vice versa when it is high. Idealized sandpile models in the mean-field limit implement in a natural way this mechanism, by means of the transport of particles through

the system up to the boundaries where they are dissipated. The content of particles regulates the activity in the system.

It is worth mentioning that going beyond the mean-field limit and turning to lattice (more realistic) systems makes things terribly complicated, and the researcher has to rely more and more on computer simulations and losses the guide of exact, or at least approximated analytical treatments. But this makes the mathematical problems that these systems pose much more interesting and exciting. For sure, researchers will devote their efforts to them for decades.

As a final point, we have to recognize that criticality and self-organized criticality are not the only ways to generate power-law distributions. In fact, much simpler processes that yield power laws exist, as reviewed in Sornette (2004), Mitzenmacher (2004), Newman (2005). A well known mechanism that escapes from the normal-distribution attractor in diffusion processes is provided by anomalous diffusion (Bouchaud and Georges 1990), and its relation with sandpiles was studied by Boguñá and Corral (1997), among others. Nevertheless, we believe the present work has clearly shown the plausibility of self-organized criticality for the explanation of earthquakes and natural hazards in general. A complementary, even more complex perspective is provided by Ben-Zion (2008).

Appendix

Properties of Power-Law Distributions

Some facts about the power-law distribution are remarkable. Let us consider the probability density $D(E) \propto 1/E^{\alpha}$, defined between $E_{\min}$ and ∞. We may first calculate its mean, i.e., the expected value of E, given by

$$\langle E \rangle = \int_{E_{\min}}^{\infty} E D(E) dE. \tag{5.87}$$

It is easy to check that, when $\alpha \leq 2$ (i.e. $b \leq 3/2$), this integral becomes infinite, so, one would say that the expected value of the energy does not exist, although we can rigurously state that that value is infinite. This second option is certainly more informative. Of course, the average energy radiated by an earthquake cannot be infinite (the Earth contains a finite amount of energy), so there is a problem extrapolating the power law up to infinity. With a normal distribution or with an exponential distribution (for example) we would not have such a problem of extrapolation, but it is worth to realize that this is a physical problem, not a mathematical problem – for instance, if instead of energy we were talking about time between some events, the mean time could perfectly be "infinite". Then, for physical reasons, there

has to be an upper limit for the validity of the Gutenberg-Richter law; however, we have no idea about how large that limit should be. In practice, the fact that the mean energy becomes infinite means that the average energy one might calculate from a series of data does not converge, no matter the number of data. Figure 5.8 illustrates this fact for the case of mean seismic moment, which is considered to be proportional to radiated energy. Summarizing, seismologists are totally ignorant about the mean energy radiated by earthquakes, due to the special properties of power-law distributions.

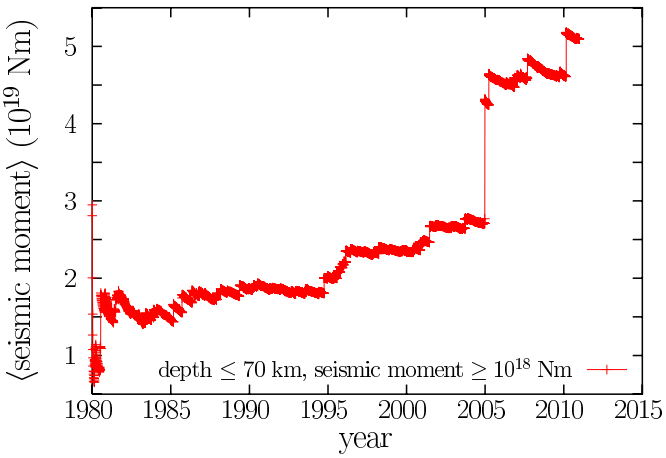


Fig. 5.8: Mean seismic moment for worldwide shallow earthquakes with seismic moment greater that 10^{18} Nm, using the CMT catalog, starting in 1980. This yields a total of 3363 events. Note that the mean value does not converge. The big jump at the end of 2004 is caused by the great Sumatra-Andaman earthquake. The radiated energy should lead to the same behavior.

Although previously we interpreted as good news the fact that most earthquakes are of small size and only very few of them are devastating, the situation is certainly not so favorable. The reason is that the rare big events, despite their scarcity, are the ones responsible for the dissipation of energy in the system. For the particular value of α we are dealing with, it is easy to check that the largest order of magnitude considered in the energy (the largest decade, or scale) contributes to the total budget more than all the other scales below. In mathematical terms,

$$\int_{E_{\min}}^{c} ED(E)dE < \int_{c}^{10c} ED(E)dE, \tag{5.88}$$

no matter how big is c (see next subsection for details).

A second peculiar property of power laws is scale invariance. Let us introduce the concept of scale transformation, considering an arbitrary function that we call $D(E)$. The idea of a scale transformation is to look at the function $D(E)$ at a different scale, as for instance, using a mathematical microscope. We can have a view of the function at the scale of meters (if E and $D(E)$ were distances) and try to see how it looks at the scale of centimeters. This is performed through a scale transformation, denoted by an operator T acting on the function $D(E)$, as

$$T[D(E)] = c_2 D(E/c_1), \tag{5.89}$$

where c_1 and c_2 are two constants called scale parameters, performing a linear transformation on E and D. In the case of the meters-centimeters example, $c_1 = c_2 = 100$.

In general, almost every function changes under a scale transformation; the exception can be found looking for the function or functions that verify the following condition,

$$D(E) = c_2 D(E/c_1). \tag{5.90}$$

It is trivial to check that a solution is given by the power-law function

$$D(E) \propto \frac{1}{E^\alpha} \tag{5.91}$$

with α given by

$$\alpha = -\frac{\ln c_2}{\ln c_1}, \tag{5.92}$$

in other words, a power law with exponent α does not change under a scale transformation if the scale factors are related through

$$c_2 = \frac{1}{c_1^\alpha} \tag{5.93}$$

Figure 5.9 shows how indeed this is the case, with $c_1 = 10$, $c_2 = \sqrt{10}$, and $D(E) = \sqrt{E}$. Note that the constant of proportionality in equation (5.91), contained in the symbol $\propto$, does not play any role here.

More importantly, it can also be demonstrated that not only the power law is a solution, but it is the only solution valid for all values of c_1 (positive real) if c_1 and c_2 are related by equation (5.93) (Takayasu 1989, Newman 2005, Christensen and Moloney 2005, Corral 2008). In summary, the condition of scale invariance demands that

$$D(E) = c_2 D(E/c_1) \text{ for all } c_1 \text{ positive real}, \tag{5.94}$$

and then, the only solution is the power law. One can verify that other solutions, as $D(E) = \sin(\ln E)$, only work for special values of c_1 and c_2.

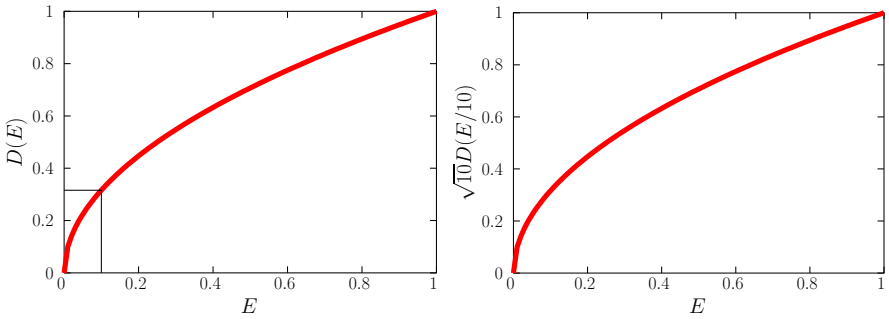


Fig. 5.9: A scale transformation acting on its corresponding scale-invariant function. The function is expanded by factors $c_1 = 10$ and $c_2 = \sqrt{10}$, in such a way that the small box at the left is the full figure at the right. The function is $D(E) = \sqrt{E}$.

Scale invariance is in fact the symmetry associated to scale transformations, in an analogous way as rotational invariance is the symmetry corresponding to rotations. If scale invariance is fulfilled, no characteristic scale can be defined for the variable E, in the same way as if there is rotational invariance in a system, this system cannot be used to point at a particular direction (a compass cannot be built from a ball). Systems do not displaying scale invariance allow to define characteristic scales, as the exponential functions defining radioactive decay lead to the definition of the unit of time in terms of the half-life.

There is, nevertheless, an important point to be taken into account here. If $D(E)$ represents a probability density (as it is the case for the energy radiated by earthquakes), then, $D(E)$ cannot be a power law for all $E \geq 0$, because it could not be normalized (its integral from 0 to ∞ would diverge). We have already mentioned that it is necessary to introduce a lower cutoff $E_{\min}$ in order to avoid this fact. Also, sometimes the power law cannot be extended to infinity, for physical reasons. So, complete scale invariance is not possible for probability distributions, and one can have only a restricted scale invariance. However, in the case of earthquakes, as both the lower limit and the upper limit are not available from observations, scale invariance plays a genuine role.

Scale invariance in the energy of earthquakes has some counter-intuitive consequences. Imagine that you arrive at a new country, and you are worried about earthquakes, and ask the people there the following question: *how big are typically earthquakes here?* Despite the innocence of such a simple question, due to scale invariance no characteristic scale for the energy can be defined and the question has no possible answer.

Dissipation of Energy in the Largest Scales

Let us consider a (continuous) power-law distribution, defined, for simplicity, between 1 and ∞, with probability density,

$$D(E) \propto \frac{1}{E^{\alpha}}. \tag{5.95}$$

We are going to see that, for a given $r > 2$ there exist values of α such that the contribution to the expected value of E from an interval $1 \leq E < c$ is always smaller than the contribution from $c \leq E < rc$, no matter how big c is.

The contribution of an interval $a \leq E < c$ to the mean value of E is

$$\int_a^c ED(E)dE \ \propto \ c^{2-\alpha} - a^{2-\alpha}. \tag{5.96}$$

Therefore,

$$\int_1^c ED(E)dE \ \propto \ c^{2-\alpha} - 1, \tag{5.97}$$

and

$$\int_c^{rc} ED(E)dE \ \propto \ c^{2-\alpha}(r^{2-\alpha} - 1). \tag{5.98}$$

In order that the last integral is larger than the previous one it is enough that

$$(r^{2-\alpha} - 1)c^{2-\alpha} > c^{2-\alpha}. \tag{5.99}$$

So, $r^{2-\alpha} > 2$ and this implies that

$$\alpha < 2 - \log_r 2. \tag{5.100}$$

For $r = 10$, the (sufficient) condition becomes $\alpha < 1.699$. In the case of earthquake radiated energy, $\alpha \simeq 1 + 2b/3 \simeq 1.667$, and equation (5.100) is fulfilled. Though, slightly larger values of α violate the condition; nevertheless, there is nothing special in taking $r = 10$ (it is not a magical number!) and we have that equation (5.100) is fulfilled for a larger r. For $r = 2$ equation (5.100) would imply $\alpha < 1$, but this is not an acceptable exponent for a power-law distribution (normalization would not be fulfilled).

Rigorous Proof of Extinction Probability

Besides graphical arguments (see Fig. 5.4), we want to provide a rigorous proof for the computation of the extinction probability in the Galton-Watson process, given by

$$P_{\text{ext}} = \lim_{t\to\infty} f^t(0), \tag{5.101}$$

where P_{ext} is properly defined only if the limit exists. To see that this is always the case, we note that $Z_t = 0 \implies Z_{t+1} = 0$. Hence, $\{Z_t = 0\} \subset \{Z_{t+1} = 0\}$ and $P(Z_t = 0) \leq P(Z_{t+1} = 0)$, so $f^t(0) \leq f^{t+1}(0)$ or, in words, (f^t) is a non-decreasing sequence. As $f([0,1]) \subset [0,1]$, we conclude that $f^t(0)$ is bounded and has a limit. To continue our proof, let us treat separately the two cases $m \leq 1$, $m > 1$. Hence,

case $m \leq 1$:

As $f(x)$ is non-convex for $x \geq 0$, it always lies above any straight line tangent to it (Spivak 1967). In particular, we consider the line tangent to $f(x)$ at the point $(1,1)$, and

$$f(x) > 1 + m(x-1) > x. \tag{5.102}$$

Hence $f(x) > x$ for $0 \leq x < 1$. Also, it is straightforward to see that $f(P_{\text{ext}}) = P_{\text{ext}}$,

$$f\left(\lim_{t\to\infty} f^t(0)\right) = \lim_{t\to\infty} f(f^t(0)) = \lim_{t\to\infty} f^{t+1}(0) = \lim_{t\to\infty} f^t(0), \tag{5.103}$$

and of course $0 \leq P_{\text{ext}} \leq 1$. So we have that $f(P_{\text{ext}}) = P_{\text{ext}}$ with $0 \leq P_{\text{ext}} \leq 1$. Summarizing, P_{ext} is a fixed point of $f(x)$ in the interval $[0,1]$, but $f(x) > x$ (strictly) in $[0,1)$. It is clear that the only option left is $P_{\text{ext}} = 1$.

case $m > 1$:

We will start showing that $P_{\text{ext}} \neq 1$ in this case. First, as already said, (f^t) is a non-decreasing sequence. Second, as $f(x)$ is continuous and $f'(1) = m > 1$, we have that $f(x) < x$ for $x \in (1-\epsilon, 1)$ for some $\epsilon > 0$. So, $f^t(0) \notin (1-\epsilon, 1)$ for all t (because it would then decrease). This means that the only way for $f^t(0)$ to have limit 1 is to "jump over" the interval $(1-\epsilon, 1)$, that is, by means of some $y < 1-\epsilon$ such that $f(y) = 1$. But such y cannot exist because then $f'(x) < 0$ at some point between y and 1.

Now we will see that the equation $f(x^*) = x^*$ has a unique solution in the interval $[0,1)$. There must be at least one solution because $f(0) > 0$, and $f(x) < x$ in $(1-\epsilon, 1)$ (here we are using Bolzano's theorem for $f(x)-x$). To see that this solution is unique, suppose there are two solutions, $0 \leq x_1 < x_2 < 1$. As we also have $f(1) = 1$, by Rolle's theorem there would exist two points y_1, y_2 such that $f'(y_1) = f'(y_2) = 1$ and $x_1 < y_1 < x_2 < y_2 < 1$, but this is impossible because $f''(x) \geq 0$ in $[0,1]$, which means that $f'(x)$ is non-decreasing and hence takes any value only once in $[0,1]$.

So, if $P_{\rm ext} \neq 1$ but $f(P_{\rm ext}) = P_{\rm ext}$, then $P_{\rm ext}$ must be the unique solution of $f(x^*) = x^*$ in $[0, 1)$.

For the sake of rigor, we must point out that some "pathological" cases would need a separate treatment, such as $f(x) = x$, but those are almost never of actual interest.

Catalan Numbers

The Catalan numbers owe their name not to a Mediterranean region but to the French-Belgian mathematician from the 19th century Eugène Charles Catalan. "His" numbers count a large variety of objects (Stanley 1999), in particular, the rooted trees that arise in the study of branching process when the number of offsprings can be $0, 1$, or 2. We can consider a tree of size s as the root (corresponding to the zero generation of the associated branching process) plus the remaining $s-1$ nodes, these latter can be distributed as a varying number of nodes associated to the first branch, $0, 1, \ldots s-1$ and the rest to the second branch, $s-1, s-2, \ldots 1, 0$, respectively. Therefore, the number of trees C_s of size s fulfills,

$$C_s = C_0 C_{s-1} + C_1 C_{s-2} + \cdots + C_{s-2} C_1 + C_{s-1} C_0, \tag{5.104}$$

where C_0 is taken equal to one, as there is only one way in which a branch can have no elements. Note that from here we obtain

$$\begin{aligned} C_1 &= (C_0)^2 = 1 \\ C_2 &= 2C_0 C_1 = 2 \\ C_3 &= 2C_0 C_2 + (C_1)^2 = 5 \\ C_4 &= 2C_3 C_0 + 2C_2 C_1 = 14 \end{aligned} \tag{5.105}$$

and so on this simple formula generates all Catalan numbers. The curious reader can check Figure 5.10, where all possible rooted trees with no more than two branches per node, of size up to 4, are shown.

If we want a closed expression for these numbers, we may define a generating function

$$h(x) = C_0 + C_1 x + C_2 x^2 + \cdots = \sum_{s=0}^{\infty} C_s x^s. \tag{5.106}$$

One can obtain an expression for $h(x)$ just using the properties of the Catalan numbers (Wilf 1994). First, let us calculate

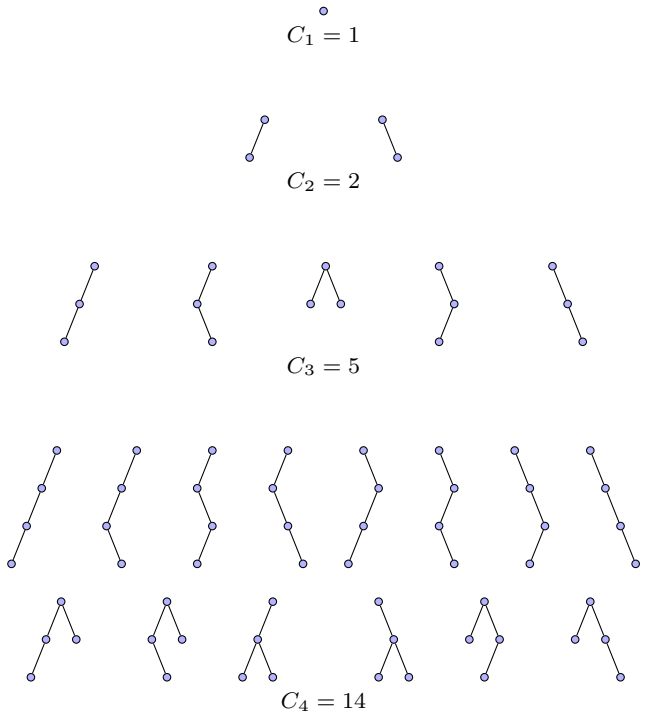


Fig. 5.10: The number of rooted trees with no more than two branches per node is shown, up to size $s = 4$. The number of such trees of a given size is given by C_s, the s-th Catalan number.

$$[h(x)]^2 = \left[\sum_{s=0}^{\infty} C_s x^s\right]^2 = \sum_{i,j=0}^{\infty} C_i C_j x^{i+j} =$$
$$= \sum_{s=0}^{\infty} \underbrace{\left[\sum_{i+j=s} C_i C_j\right]}_{C_{s+1}} x^s = \frac{1}{x}\sum_{s=0}^{\infty} C_{s+1} x^{s+1} = \frac{h(x) - C_0}{x}$$

As we know that $C_0 = 1$, we end up with a quadratic equation for $h(x)$, namely,

$$x[h(x)]^2 - h(x) + 1 = 0, \tag{5.107}$$

which allows us to isolate $h(x)$,

$$h(x) = \frac{1 \pm \sqrt{1 - 4x}}{2x}. \tag{5.108}$$

One of both functions (depending on the $\pm$ sign) is then the generating function of the Catalan numbers. We are going to recover these numbers from its generating function. First, one needs the Taylor expansion of $\sqrt{1-x}$ around $x=0$, which is

$$\sqrt{1-x} = 1 - \frac{x}{2} - \frac{1}{4}\frac{x^2}{2!} - \frac{3}{8}\frac{x^3}{3!} - \cdots = 1 - \frac{x}{2} - \sum_{s=1}^{\infty} \frac{(2s-1)!!}{2^{s+1}(s+1)!} x^{s+1}, \quad (5.109)$$

where, remember, $n!! = n(n-2)\cdots 1$, and so,

$$\sqrt{1-4x} = 1 - 2x - \sum_{s=1}^{\infty} \frac{(2s-1)!!2^{s+1}}{(s+1)!} x^{s+1}. \quad (5.110)$$

Then, substituting in $h(x)$, one can realize that only the minus sign can correspond to a generating function, and

$$h(x) = 1 + \frac{1}{2x} \sum_{s=1}^{\infty} \frac{(2s-1)!!2^{s+1}}{(s+1)!} x^{s+1} = 1 + \sum_{s=1}^{\infty} \frac{(2s-1)!!2^{s}}{(s+1)!} x^{s}, \quad (5.111)$$

from where we obtain a first expression for the Catalan numbers,

$$C_s = \frac{(2s-1)!!2^s}{(s+1)!} \text{ for } s \geq 1. \quad (5.112)$$

A more comfortable formula can be obtained using that

$$(2s)! = (2s)!!(2s-1)!! = s!2^s(2s-1)!!, \quad (5.113)$$

and then one finds,

$$C_s = \frac{(2s)!}{s!(s+1)!} = \frac{1}{s+1}\binom{2s}{s}, \quad (5.114)$$

the standard expression for the Catalan numbers, now valid for all $s \geq 0$.

Normalization and non-normalization of the total size distribution

We are going to illustrate how the total size probability distribution, $P(S = s)$, is only normalized in the subcritical and critical cases. We use the binomial distribution for the distribution of the number of offsprings, with $k = 0, 1$ and 2. From the main text, we know that

$$P(S=s)=\frac{1}{s+1}\binom{2s}{s}p^{s-1}(1-p)^{s+1} \text{ with } s=1,2,\ldots \tag{5.115}$$

It can be checked, using the generating function of the Catalan numbers, that this expression is normalized for $p \leq 1/2$ but not for $p > 1/2$. In order to see this, let us first consider the generating function of the Catalan numbers, derived in the previous subsection of the Appendix,

$$h(x)=\sum_{s=0}^{\infty} C_s x^s = \frac{1-\sqrt{1-4x}}{2x}. \tag{5.116}$$

Then, introducing $q = 1 - p$,

$$\sum_{s=1}^{\infty} P(S=s) = \frac{q}{p}\sum_{s=1}^{\infty} C_s (pq)^s = \frac{q}{p}\left(h(pq)-1\right), \tag{5.117}$$

and using the expression for $h(x)$,

$$h(pq)=\frac{1-\sqrt{1-4pq}}{2pq}=\frac{1-\sqrt{(1-2p)^2}}{2pq}=\frac{1-|1-2p|}{2pq}. \tag{5.118}$$

We can distinguish two cases, first, $p \leq 1/2$, for which,

$$h(pq)-1=\frac{1}{q}-1=\frac{p}{q}=\frac{\min(p,q)}{\max(p,q)}, \tag{5.119}$$

and for the opposite case, $p \geq 1/2$,

$$h(pq)-1=\frac{1}{p}-1=\frac{q}{p}=\frac{\min(p,q)}{\max(p,q)}. \tag{5.120}$$

Therefore,

$$\sum_{s=1}^{\infty} P(S=s) = \frac{q}{p}\frac{\min(p,q)}{\max(p,q)} = \begin{cases} 1 & \text{for } p \leq 1/2 \\ \left(\frac{q}{p}\right)^2 & \text{for } p \geq 1/2 \end{cases} \tag{5.121}$$

Remembering the results for the extinction probability for the binomial distribution,

$$\sum_{s=1}^{\infty} P(S=s) = P_{ext}, \tag{5.122}$$

which obviously is not normalized for $p > 1/2$. We could also have arrived to the same result using, not the generating function of the Catalan numbers, but the generating function $g(x)$ of the size S.

Stirling's Approximation

Usually, Stirling's formula is demonstrated by means of the Euler-Maclaurin formula. However, if one knows some elementary properties of the gamma distribution, Stirling's formula arises almost spontaneously, by means of a probabilistic trick.

Remember that the factorial is associated to the gamma function, $n! = \Gamma(n+1)$, which is defined as

$$\Gamma(\gamma) = \int_0^\infty y^{\gamma-1} e^{-y} dy \tag{5.123}$$

for $\gamma > 0$ (Abramowitz and Stegun 1965). This allows to introduce the gamma distribution (Durrett 2010), with probability density given by

$$\frac{1}{\Gamma(\gamma)} y^{\gamma-1} e^{-y} \tag{5.124}$$

for $y \geq 0$ (and zero otherwise), and with mean γ and variance γ.

It turns out that the gamma distribution arises as a sum of a number γ of independent exponential random variables, each with density e^{-y} (this can be easily demonstrated through successive convolutions of the exponentials, see Durrett (2010)). But using the central limit theorem, the gamma distribution will converge, in the limit $\gamma \to \infty$, to a normal distribution (see Fig. 5.11), with mean μ and standard deviation σ (in this case the notation is different to the rest of the chapter).

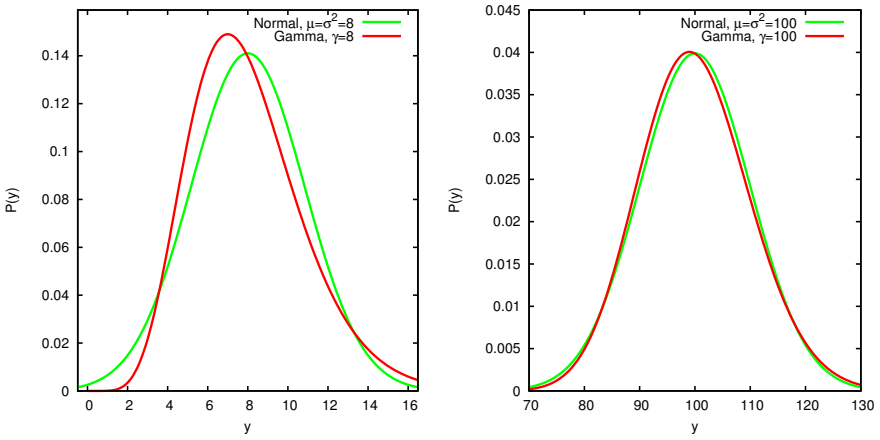


Fig. 5.11: Approaching of the normal distribution by the gamma distribution, adding 8 and 100 exponentials, respectively. The central limit theorem allows the derivation of Stirling's approximation.

Then, it will be possible to transform the gamma function into a Gaussian integral. Indeed,

$$n! = \Gamma(n+1) = \int_0^\infty y^n e^{-y} dy \to C \int_0^\infty \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy. \tag{5.125}$$

The key point is to find the value of C for which both functions overlap. This happens around the mean or the mode of both distributions, corresponding, respectively, to $y = \gamma = n + 1 \simeq n$ and $y = \mu$. Substituting both values in

$$y^n e^{-y} = C \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) \tag{5.126}$$

we get

$$C = \left(\frac{n}{e}\right)^n \tag{5.127}$$

and therefore, looking for the normal probability density inside the integral,

$$n! = \Gamma(n+1) \to \sqrt{2\pi}\sigma C \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy. \tag{5.128}$$

The value of σ is obtained from $\sigma^2 = \gamma = n + 1$ (for independent random variables the variance of a sum is the sum of variances, which is one for each exponential distribution in our sum). Substituting, and replacing the lower integration limit by $-\infty$, due to the fact that the standard deviation $\sigma \simeq \sqrt{n}$ is much smaller than the mean $\mu \simeq n$, one obtains,

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n, \tag{5.129}$$

valid, remember, in the limit $n \to \infty$. This proof has some parts in common with the more elaborated one of Khan (1974) and less resemblance with that of van den Berg (1995).

Acknowledgements We would like to dedicate this work to the colorful scientist Per Bak, in the 25 years of his invention of self-organized criticality and in the 10th anniversary of his untimely death. The chapter originates, in part, from a lecture that one of the authors gave at the 2011 Fall Meeting of the *American Geophysical Union*. In this regard, we thank Armin Bunde, and also Tom Davis, for making his notes on the Catalan numbers publicly available on the Internet, and Anna Deluca and Gunnar Pruessner, for discussions. Cecília M. Clos provided valuable graphical-design assistance. Funding has come from Spanish projects FIS2009-09508 and 2009-SGR-164.

References

Abramowitz, M. and Stegun, I. A., editors (1965). *Handbook of Mathematical Functions*. Dover, New York.

Alligood, K., Sauer, T., and Yorke, J. (1997). *Chaos. An Introduction to Dynamical Systems*. Textbooks in Mathematical Sciences. Springer-Verlag.

Alstrøm, P. (1988). Mean-field exponents for self-organized critical phenomena. 38:4905–4906.

Åström, J., Stefano, P. C. F. D., Pröbst, F., Stodolsky, L., Timonen, J., Bucci, C., Cooper, S., Cozzini, C., Feilitzsch, F., Kraus, H., Marchese, J., Meier, O., Nagel, U., Ramachers, Y., Seidel, W., Sisti, M., Uchaikin, S., and Zerle, L. (2006). Fracture processes observed with a cryogenic detector. *Phys. Lett. A*, 356:262–266.

Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. Copernicus, New York.

Bak, P., Christensen, K., Danon, L., and Scanlon, T. (2002). Unified scaling law for earthquakes. *Phys. Rev. Lett.*, 88:178501.

Bak, P. and Sneppen, K. (1993). Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.*, 24:4083–4086.

Bak, P. and Tang, C. (1989). Earthquakes as a self-organized critical phenomenon. *J. Geophys. Res.*, 94:15635–15637,.

Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: an explanation of $1/f$ noise. *Phys. Rev. Lett.*, 59:381–384.

Ben-Zion, Y. (2008). Collective behavior of earthquakes and faults: continuum-discrete transitions, progressive evolutionary changes, and different dynamic regimes. *Rev. Geophys.*, 46:RG4006.

Boguñá, M. and Corral, A. (1997). Long-tailed trapping times and Lévy flights in a self-organized critical granular system. *Phys. Rev. Lett.*, 78:4950–4953.

Bouchaud, J.-P. and Georges, A. (1990). Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications. *Phys. Rep.*, 195:127–293.

Burridge, R. and Knopoff, L. (1967). Model and theoretical seismicity. *Bull. Seismol. Soc. Am.*, 57:341–371.

Burroughs, S. M. and Tebbens, S. F. (2005). Power-law scaling and probabilistic forecasting of tsunami runup heights. *Pure Appl. Geophys.*, 162:331–342.

Christensen, K., Corral, A., Frette, V., Feder, J., and Jøssang, T. (1996). Tracer dispersion in a self-organized critical system. *Phys. Rev. Lett.*, 77:107–110.

Christensen, K. and Moloney, N. R. (2005). *Complexity and Criticality*. Imperial College Press, London.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Rev.*, 51:661–703.

Corral, A. (2004a). Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Phys. Rev. Lett.*, 92:108501.

Corral, A. (2004b). Universal local versus unified global scaling laws in the statistics of seismicity. *Physica A*, 340:590–597.

Corral, A. (2008). Scaling and universality in the dynamics of seismic occurrence and beyond. In Carpinteri, A. and Lacidogna, G., editors, *Acoustic Emission and Critical Phenomena*, pages 225–244. Taylor and Francis, London.

Corral, A. (2010). Tropical cyclones as a critical phenomenon. In Elsner, J. B., Hodges, R. E., Malmstadt, J. C., and Scheitlin, K. N., editors, *Hurricanes and Climate Change: Volume 2*, pages 81–99. Springer, Heidelberg.

Corral, A., Ossó, A., and Llebot, J. E. (2010). Scaling of tropical-cyclone dissipation. *Nature Phys.*, 6:693–696.

Durrett, R. (2010). *Probability: Theory and Examples*. Cambridge University Press, Cambridge, 4th edition.

Freeman, M. P. and Watkins, N. W. (2002). The heavens in a pile of sand. *Science*, 298:979–980.

Grimmett, G. and Stirzaker, D. (2001). *Probability and Random Processes*. Oxford University Press, Oxford, 3rd edition.

Gutenberg, B. and Richter, C. F. (1944). Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.*, 34:185–188.

Harris, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.

Hawkins, D. and Ulam, S. (1944). Theory of multiplicative processes I. *LANL report*, LA - 171.

Helmstetter, A. and Sornette, D. (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *J. Geophys. Res.*, 107 B:2237.

Hergarten, S. (2012). Branching with local probability as a paradigm of self-organized criticality. *Phys. Rev. Lett.*, 109:148001.

Ito, K. (1995). Punctuated-equilibrium model of biological evolution is also a self-organized-criticality model of earthquakes. *Phys. Rev. E*, 52:3232–3233.

Ito, K. and Matsuzaki, M. (1990). Earthquakes as self-organized critical phenomena. *J. Geophys. Res.*, 95:6853–6860.

Jensen, H. J. (1998). *Self-Organized Criticality*. Cambridge University Press, Cambridge.

Kadanoff, L. P. (1991). Complex structures from simple systems. *Phys. Today*, 44:9–11.

Kagan, Y. Y. (2010). Earthquake size distribution: Power-law with exponent $\beta \equiv 1/2$? *Tectonophys.*, 490:103–114.

Kagan, Y. Y., Jackson, D. D., and Geller, R. J. (2012). Characteristic earthquake model, 1884 - 2011, r.i.p. *ArXiv*, 1207:4836.

Kanamori, H. and Brodsky, E. E. (2004). The physics of earthquakes. *Rep. Prog. Phys.*, 67:1429–1496.

Kanamori, H. and Mori, J. (2000). Microscopic processes on a fault plane and their implications for earthquake dynamics. In Boschi, E., Ekström, G., and Morelli, A., editors, *Problems in geophysics for the new millennium*, pages 73–88. Editrice Compositori, Bolonia.

Kendall, D. G. (1966). Branching processes since 1873. *J. London Math. Soc.*, 41:385–406.

Kendall, D. G. (1975). The genealogy of genealogy branching processes before (and after) 1873. *Bull. London Math. Soc.*, 7:225–253.

Khan, R. A. (1974). A probabilistic proof of Stirling's formula. *Am. Math. Mon.*, 81(4):366–369.

Kwiatek, G., Plenkers, K., Nakatani, M., Yabe, Y., Dresen, G., and JAGUARS-Group (2010). Frequency-magnitude characteristics down to magnitude -4.4 for induced seismicity recorded at Mponeng gold mine, South Africa. *Bull. Seismol. Soc. Am.*, 100:1165–1173.

Lahaie, F. and Grasso, J. R. (1998). A fluid-rock interaction cellular automaton of volcano mechanics: Application to the Piton de la Fournaise. *J. Geophys. Res.*, 103 B:9637–9650.

Main, I. (1996). Statistical physics, seismogenesis, and seismic hazard. *Rev. Geophys.*, 34:433–462.

Malamud, B. D. (2004). Tails of natural hazards. *Phys. World*, 17 (8):31–35.

Malamud, B. D., Millington, J. D. A., and Perry, G. L. W. (2005). Characterizing wildfire regimes in the United States. *Proc. Natl. Acad. Sci. USA*, 102:4694–4699.

Manna, S. S. (1991). Two-state model of self-organized criticality. *J. Phys. A: Math Gen.*, 24:L363–L369.

Mantegna, R. N. and Stanley, H. E. (1999). *An Introduction to Econophysics*. Cambridge University Press, Cambridge.

McClelland, L., Simkin, T., Summers, M., Nielsen, E., and Stein, T. C., editors (1989). *Global volcanism 1975-1985*. Prentice Hall, Englewood Cliffs.

Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Math.*, 1 (2):226–251.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Cont. Phys.*, 46:323 –351.

Newman, M. E. J. (2011). Resource letter CS-1: Complex systems. *Am. J. Phys.*, 79:800–810.

Ogata, Y. (1999). Seismicity analysis through point-process modeling: a review. *Pure Appl. Geophys.*, 155:471–507.

Olami, Z., Feder, H. J. S., and Christensen, K. (1992). Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Phys. Rev. Lett.*, 68:1244–1247.

Otsuka, M. (1971). A simulation of earthquake occurrence part 2. magnitude-frequency relation of earthquakes (in Japanese). *Zisin*, 24:215–227.

Otsuka, M. (1972). A chain-reaction-type source model as a tool to interpret the magnitude-frequency relation of earthquakes. *J. Phys. Earth*, 20:35–45.

Otter, R. (1949). The multiplicative process. *Ann. Math. Stat.*, 20:206–224.

Peters, O., Deluca, A., Corral, A., Neelin, J. D., and Holloway, C. E. (2010). Universality of rain event size distributions. *J. Stat. Mech.*, P11030.

Peters, O. and Neelin, J. D. (2006). Critical phenomena in atmospheric precipitation. *Nature Phys.*, 2:393–396.

Pruessner, G. and Peters, O. (2006). Self-organized criticality and absorbing states: Lessons from the Ising model. *Phys. Rev. E*, 73:025106(R).

Ross, S. M. (2002). *A First Course in Probability*. Prentice Hall, Englewood Cliffs.

Sornette, A. and Sornette, D. (1989). Self-organized criticality and earthquakes. *Europhys. Lett.*, 9:197–202.

Sornette, D. (1992). Critical phase transitions made self-organized: a dynamical system feedback mechanism for self-organized criticality. *J. Phys. I France*, 2(11):2065–2073.

Sornette, D. (2004). *Critical Phenomena in Natural Sciences*. Springer, Berlin, 2nd edition.

Spivak, M. (1967). *Calculus*. Publish or Perish, Inc.

Stanley, R. P. (1999). *Enumerative Combinatorics, Volume 2*. Cambridge University Press, Cambridge.

Stein, R. S. (2002). Parkfield's unfulfilled promise. *Nature*, 419:257–258.

Takayasu, H. (1989). *Fractals in the Physical Sciences*. Manchester University Press, Manchester.

Ulam, S. M. (1990). *Analogies Between Analogies: The Mathematical Reports of S.M. Ulam and his Los Alamos Collaborators*. University of California Press.

Ulam, S. M. (1991). *Adventures of a mathematician*. University of California Press.

Utsu, T. (1999). Representation and analysis of earthquake size distribution: a historical review and some new approaches. *Pure Appl. Geophys.*, 155:509–535.

van den Berg, I. P. (1995). External probability. In Diener, F. and Diener, M., editors, *Nonstandard Analysis in Practice*, page 171. Springer-Verlag, Berlin.

Vere-Jones, D. (1976). A branching model for crack propagation. *Pure Appl. Geophys.*, 114:711–725.

Watson, H. W. and Galton, F. (1875). On the probability of the extinction of families. *J. Anthropol. Inst. Great Britain Ireland*, 4:138–144.

Wilf, H. S. (1994). *generatingfunctionology*. Academic Press.

Zapperi, S., Lauritsen, K. B., and Stanley, H. E. (1995). Self-organized branching processes: Mean-field theory for avalanches. *Phys. Rev. Lett.*, 75:4071–4074.

Chapter 6
Power Laws of Recurrence Networks*

Yong Zou, Jobst Heitzig, Jürgen Kurths

Abstract

Recurrence networks are a novel tool of nonlinear time series analysis allowing the characterisation of higher-order geometric properties of complex dynamical systems based on recurrences in phase space, which are a fundamental concept in classical mechanics. In this Chapter, we demonstrate that recurrence networks obtained from various deterministic model systems as well as experimental data naturally display power-law degree distributions with scaling exponents γ that can be derived exclusively from the systems' invariant densities. For one-dimensional maps, we show analytically that γ is not related to the fractal dimension. For continuous systems, we find either power-laws with an exponent γ depending on a suitable notion of local dimension, or with fixed $\gamma = 1$. Some computational techniques relevant to the construction of recurrence networks have been discussed in detail.

6.1 Introduction

Power-law distributions have been widely observed in diverse fields such as seismology, economy, ecology, and finance in the context of critical phenomena Gutenberg (1944), Marquet (2005), Mantegna (1995), Farmer (2004). In

Department of Physics, East China Normal University, Shanghai, 200062, China · Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong · Potsdam Institute for Climate Impact Research, P. O. Box 601203, 14412 Potsdam, Germany · Department of Physics, Humboldt University Berlin, Newtonstr. 15, 12489 Berlin, Germany · Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB243UE, United Kingdom

* The main part of this chapter is compiled from the publication Zou (2012).

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

many cases, the underlying complex systems can be regarded as networks of mutually interacting subsystems with a complex structural organisation. Specifically, numerous examples have been found for hierarchical structures in the connectivity of such complex networks, *i.e.,* the presence of scale-free distributions $P(k) \sim k^{-\gamma}$ of the node degrees Albert (2002), Newman (2003). Such hierarchical organisation is particularly well expressed in network of networks, or interdependent networks, which constitute an emerging and important new field of complex network research Gao (2012), Donges (b). The interrelationships between the non-trivial structural properties of complex networks and the resulting dynamics of the mutually interacting subsystems are subject of intensive research Boccaletti (2006), Arenas (2008).

Among other developments, one of the main recent achievements of complex network theory are various conceptionally different approaches for statistically characterising dynamical systems by graph-theoretical methods Zhang (2006), Lacasa (2008), Xu (2008), Marwan (2009). In this Chapter, we report and thoroughly explain the emergence of power-laws in the degree distribution of so-called *recurrence networks* (RNs) Marwan (2009), Donner (b,a), Donges (a) for various paradigmatic model systems as well as experimental data. RNs encode the underlying system's *recurrences in phase space* and are based on a fundamental concept in classical physics Poincaré (1890). Due to their direct link to dynamical systems theory, RNs are probably the most widely applicable type of complex networks inferred from time series introduced so far. Although the system's temporal evolution cannot be reconstructed from the RN, this representation allows for an analysis of the attractor's geometry in phase space using techniques from network theory. Specifically, nodes represent individual state vectors, and pairs of nodes are linked when they are mutually closer than some threshold distance $\varepsilon > 0$ Marwan (2007) (cf. Fig. 6.1), which is a key parameter of this method Donner (b). Thus, the adjacency matrix of a RN reads

$$A_{ij} = \Theta(\varepsilon - \|\boldsymbol{x}_i - \boldsymbol{x}_j\|) - \delta_{ij}, \tag{6.1}$$

where $\Theta(\cdot)$ is the Heaviside function, δ_{ij} Kronecker's delta, and $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ the coordinates of a pair of state vectors (represented by RN nodes) in phase space. According to this definition, RNs are random geometric graphs Hermann (2003) (*i.e.*, undirected *spatial networks* Bartélemy (2011)), where the spatial distribution of nodes is completely determined by the probability density function of the invariant measure of the dynamical system under study, and links are established according to the distance in phase space. The *degree* of a node i is defined as $k_i = \sum_{j=1}^{N} A_{ij}$. Therefore the degree distribution $P(k)$ of the associated network is defined to be the fraction of nodes in the network with degree k. If there are N nodes in total in a network and n_k of them have degree k, we have $P(k) = n_k/N$. Consequently, their degree distribution $P(k)$ directly relates to the system's invariant density $p(x)$.

In this work, we demonstrate the emergence of scaling in the degree distributions of RNs and provide some evidence that this phenomenon is (unlike many other scaling exponents occurring in the context of dynamical systems) commonly unrelated to the fractal attractor dimension, except for some interesting special cases. Instead, the power-laws naturally arise from the variability of the invariant density $p(x)$ of the system (*i.e.,* peaks or singularities of p), as we will show numerically as well as explain theoretically.

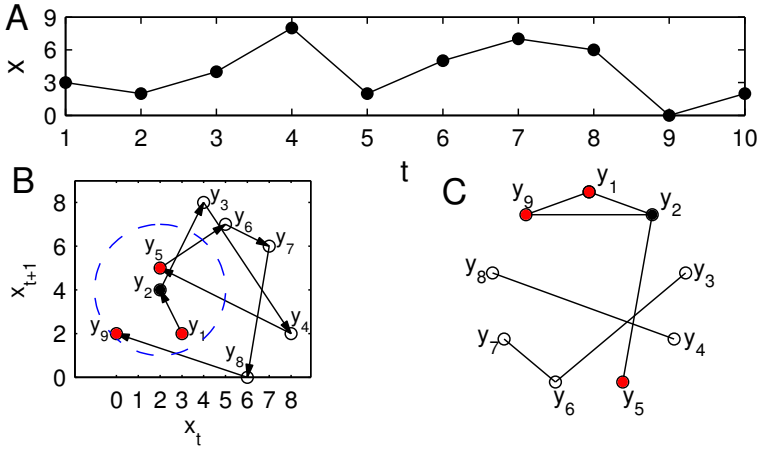


Fig. 6.1: (Colour online) Construction of a RN: (A) time series x_t, (B) reconstructed two-dimensional phase space trajectory with state vectors $\mathbf{y}_t = (x_t, x_{t+1})$ and an exemplary Euclidean ε-ball around y_2 (dashed) for $\varepsilon = 3$. (C) Resulting recurrence network for the specific chosen ε. Note that the arrow lines of (B) correspond to the temporal direction, while lines in (C) are links of the network.

6.2 Power-law scaling and singularities of the invariant density

As initial examples, Fig. 6.2 illustrates the presence of power-law degree distributions in the RNs obtained for several prototypical low-dimensional chaotic systems with a suitable choice of the systems' characteristic parameters: (i) the Rössler system in spiral-chaos regime: $\dot{x} = -y - z$, $\dot{y} = x + 0.2y$, $\dot{z} = 0.2 + z(x - 5.7)$; (ii) the Lorenz system: $\dot{x} = 10(y - x)$, $\dot{y} = x(r - z) - y$, $\dot{z} = xy - 8/3z$; and (iii) the Hénon map: $x_{n+1} = 1 - 1.4x_n^2 + y_n$, $y_{n+1} = 0.3x_n$. For the Rössler and Lorenz systems, a proper time discretisation has been used as explained in detail in the figure caption.

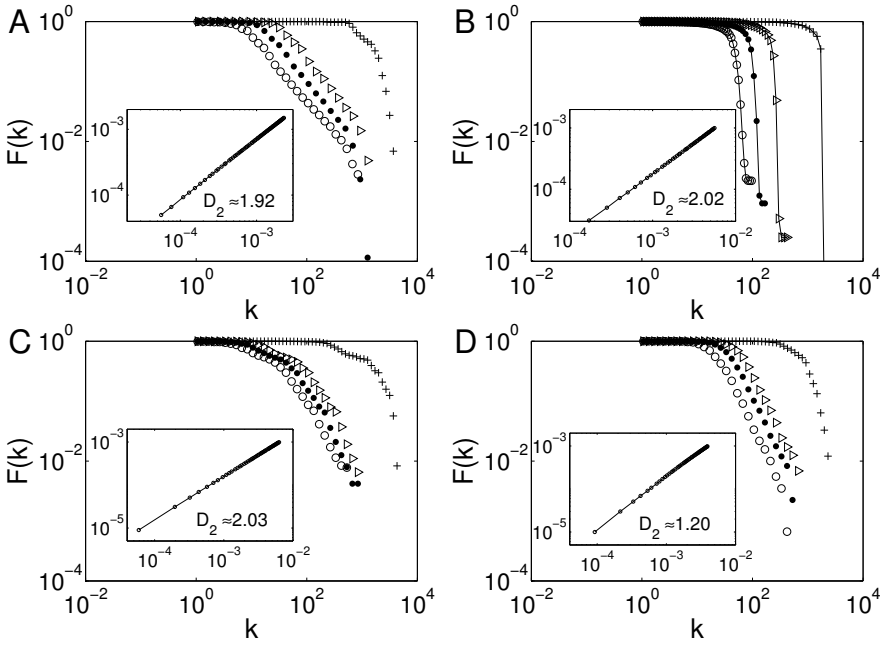


Fig. 6.2: Cumulative degree distributions $F(k) = \sum_{k'=k}^{\infty} P(k')$ of the RNs obtained for several discrete maps: (A) x-component of the first return map of the Rössler system ($\gamma = 2.16 \pm 0.03$), (B) map of consecutive local maxima of the z-component of the Lorenz system with $r = 28$ (no scaling), (C) as in (B) with $r = 90$ ($\gamma = 2.64 \pm 0.18$), and (D) Hénon map ($\gamma = 2.88 \pm 0.04$). Estimates of $\gamma > 1$ have been obtained from the cumulative degree distributions $F(k) \sim k^{1-\gamma}$ (in the following figures, we show either $F(k)$ or $P(k)$ when sufficiently straight in a log-log plot) by means of a maximum likelihood approach Clauset (2009) as averages considering 100 different values of ρ for which a power-law appears. For each ρ, 5 different realisations with random initial conditions are used. We consider RNs of size $N = 2 \times 10^5$ and the Euclidean norm in all examples discussed in this Chapter: Four cases were chosen for illustration corresponding to a link density of $\rho_1 = 0.02\%$ ($\circ$), $\rho_2 = 0.03\%$ ($\bullet$), $\rho_3 = 0.05\%$ ($\triangleright$), and $\rho_4 = 3\%$ ($+$). A power-law is hardly detectable for ρ_4. Insets: log-log plot of the correlation sum $C(\varepsilon)$ vs. ε, where the correlation dimension D_2 is estimated by linear regression and agrees well with values from the literature Grassberger (1983), Sprott (2003), Kantz (1997).

In all cases, scaling emerges only if the distance threshold ε is chosen small enough, which corresponds to a small average degree $\langle k \rangle$ and link density $\rho = \langle k \rangle / (N-1)$ of the resulting network (N being the number of nodes). We note that the respective range of ε should be sufficiently higher than the threshold for which a giant component exists Donges (2012). The size of the scaling regime decreases with growing ρ and becomes hardly detectable for $\rho \gtrsim 1\%$ (Fig. 6.2). The distance threshold ε also occurs in dimension estimation where the limit $\varepsilon \to 0$ is taken (*e.g.,* Farmer (1983)). In contrast,

a RN is based on one finite ε. Our smallest ε are, however, still large enough to avoid the problem of lack of neighbours Sprott (2003) since the correlation integral $C(\varepsilon)$ still shows the same dependency on ε as for larger values (see insets in Fig. 6.2).

6.2.1 One-dimensional maps: Analytical theory

In order to obtain an analytical explanation of the emergence of power-laws in the degree distribution $P(k)$ of a RN, we consider recent findings from the more general theory of random spatial networks Herrmann (2003). In this framework, nodes are randomly sampled from some given spatial probability density function $p(x)$. In the special case of a RN, we identify this space with the phase space of a (dissipative) dynamical system and the nodes of the network with individual states sampled at discrete times. Furthermore, we make the following assumptions: (i) The system under study is ergodic. (ii) The sampled trajectory is already close to its attractor (i.e., we exclude the presence of transient behavior). (iii) The sampling times are generic, i.e., the sampling interval is co-prime to any possible periods of the system. Under the validity of these assumptions, the nodes can be interpreted as being sampled from the probability density function $p(x)$ of the invariant measure μ of the attractor Eckmann (1985), for short refered to as the invariant density in the following.

For a general spatial network, the degree distribution $P(k)$ can be derived from $p(x)$ in the limit of large network size N as

$$P(k) = \int dx\, p(x) e^{-\alpha p(x)} (\alpha p(x))^k / k! \tag{6.2}$$

with $\alpha = \langle k \rangle \,/ \int dx\, p(x)^2$ (note that the computation of $\langle k \rangle$ involves integration of $p(x)$ over the ε-neighbourhood of all points x and thus implicitly depends on the specifically chosen ε as well as the sample size N). In order to understand this relationship, note that for each x, the probability that a sampled point falls into the ε-ball centered at x is proportional to $p(x)$. Hence, the degree of a node at x has a binomial distribution. For sufficiently large N, the latter can be approximated by a Poissonian distribution with parameter $\alpha p(x)$, leading to Eq. (6.2).

For the special case of a one-dimensional phase space, we can use the general expression (6.2) in order to derive an explicit criterion for the emergence of scale-free degree distributions of the RN obtained from dynamical systems. For this purpose, let us assume that the invariant density $p(x)$ on the attractor fulfils the following weak condition: For almost every y, there are only finitely many different x with $p(x) = y/\alpha$ (i.e., the set of values y for which there are infinitely many x with $p(x) = y/\alpha$ has zero Lebesgue-

measure). Given the validity of this assumption, we can simplify Eq. (6.2) by substituting $y = \alpha p(x)$: Putting

$$g(y) = \sum_{x \in p^{-1}(y/\alpha)} |p'(x)|^{-1}, \tag{6.3}$$

we obtain

$$P(k) = \int_0^\infty dy \, \frac{e^{-y} y^{k+1}}{\alpha k!} g(y). \tag{6.4}$$

The function g describes the distribution of the values of the function p. Note that

$$q_k(y) = e^{-y} y^{k+1}/(k+1)! \tag{6.5}$$

is the density of a gamma distribution with scale parameter one and shape parameter $k+2$, so that

$$E_{q_k}(g) = \int_0^\infty dy \, q_k(y) g(y) \tag{6.6}$$

is the expected valued of $g(y)$ when y comes from that gamma distribution. Hence

$$P(k) = \frac{k+1}{\alpha} E_{q_k}(g). \tag{6.7}$$

The bulk of the gamma distribution $q_k(y)$ resides almost symmetrically in the interval $J_k = k + 1 \pm 2\sqrt{k+2}$, so if g has low curvature within this interval, then its expectation $E_{q_k}(g)$ is approximately equal to its value at $y = k+1$, and

$$P(k) \approx \frac{k+1}{\alpha} g(k+1) = \frac{k+1}{\alpha} \sum_{x \in p^{-1}(\frac{k+1}{\alpha})} |p'(x)|^{-1}. \tag{6.8}$$

Now, if either $g(y)$ or $p(x)$ follows a power-law, so will $P(k)$: Assuming $g(y) \sim y^{-1-\gamma}$ for some $\gamma > 0$, Eq. (6.8) gives $P(k) \sim k^{-\gamma}$, and this approximation is good since the relative error in the linear Taylor approximation of $g(y)$ inside the interval J_k about its center $y = k+1$ is only of the order of $\sqrt{k}^2 k^{-3-\gamma}/k^{-1-\gamma} = k^{-1}$. In particular, a power-law in the invariant density of the form $p(x) \sim |x - x_0|^{-1/\gamma}$ for x close to x_0 and with $\gamma \neq 0$ will imply $g(y) \sim y^{-1-\gamma}$ and thus $P(k) \sim k^{-\gamma}$. From these considerations, we can infer that the slower the invariant density decays, the faster does the scale-free degree distribution of the associated RN. This implies that if $p(x)$ has a power-law-shaped peak at some state x_0, *i.e.,* $p(x) \sim |x - x_0|^{-1/\gamma}$ for some $\gamma > 0$, the degree distribution $P(k)$ also follows a power-law but with the reciprocal exponent, $P(k) \sim k^{-\gamma}$.

More generally, we can deduce that the presence of singularities in the invariant density is the key feature determining whether or not the resulting RN has a power-law degree distribution. This relationship can be intuitively understood: If $p(x)$ has a singularity at some point x_0 in phase space, then

a time series of the associated dynamical system will return very often to the neighbourhood of x_0. Hence, nodes with a high degree will accumulate close to the singularity. If the resulting invariant density obeys a power-law decay, Eq. (6.2) implies the emergence of a power-law degree distribution. If there is more than one singularity of $p(x)$, one can expect the resulting degree distribution being related to a weighted sum of the influences of these points. Vice versa, the presence of a power-law degree distribution in the RN of a dynamical system requires the existence of a power-law in the invariant density, *i.e.,* the presence of a singularity.

Note that not all invariant densities lead to power-laws: if $p(x)$ is Gaussian, $g(y)$ is again smooth enough to use Eq. (6.8). However, in this case, Eq. (6.8) does *not* indicate a power-law: $P(k) \approx 2(-2\ln k\sqrt{2\pi}/\alpha)^{-1/2}$. We therefore conjecture that a local power-law in the invariant density is a necessary and sufficient condition for the emergence of a scale-free RN. Beyond the explicit results for one-dimensional systems as discussed above, we further show below that the emergence of a power-law scaling is also possible in higher dimensions and conjecture that this requires the presence of a dynamically invariant object (*e.g.,* an unstable or hyperbolic fixed point) close to which the invariant density scales as a power-law at least in one direction.

6.2.2 Example: Generalized logistic map

As a widely studied example for a discrete-time system (map), in the following we study the properties of the *generalized logistic map* or β-map Lyra (1998)

$$f(x) = 1 - |2x - 1|^{\beta} \tag{6.9}$$

with $\beta \geqslant \frac{1}{2}$ in some detail. For the special cases $\beta = \frac{1}{2}, 1, 2$, this map corresponds to the cusp map, tent map, and standard logistic map, respectively. In general, (6.9) maps the unit interval $[0, 1]$ onto itself and is symmetric with a maximum of 1 at $x = \frac{1}{2}$, thus having two pre-images for each $x < 1$.

In order to analytically derive an expression for the degree distribution of RNs associated with the generalized logistic map, we start by computing the absolute value of the derivative of $f(x)$, which reads $|f'(x)| = 2\beta|2x - 1|^{\beta-1}$. Because of the conservation of probability, the invariant density $p(x)$ must fulfil

$$p(x) + p(1 - x) = 2\beta|2x - 1|^{\beta-1}p(1 - |2x - 1|^{\beta}). \tag{6.10}$$

For the case $\beta > \frac{1}{2}$, instead of trying to solve this exactly, we focus on its behavior at $x = 0$ and $x = 1$ and near the point $x = 1/2$ that is mapped to $x = 1$. First consider that $x = (1 + \delta^{1/\beta})/2$ with $0 < \delta \ll 1$. Then Eq. (6.10) reads $2p(0.5) \approx 2\beta\delta^{(\beta-1)/\beta}p(1 - \delta)$. Hence the density close to the right peak has a power-law of the form

$$p(1-\delta) \approx p(0.5)\delta^{(1-\beta)/\beta}/\beta. \tag{6.11}$$

Now consider that $x = \delta$ with $0 < \delta \ll 1$. Then Eq. (6.10) reads $p(\delta) + p(1-\delta) \approx 2\beta p(2\beta\delta)$. Combining both approximations, we get

$$2\beta p(2\beta\delta) \approx p(\delta) + p(0.5)\delta^{(1-\beta)/\beta}/\beta. \tag{6.12}$$

Let us next check whether (i) $p(\delta) = o(\delta^{(1-\beta)/\beta})$, (ii) $\delta^{(1-\beta)/\beta} = o(p(\delta))$, or (iii) $p(\delta) \sim \delta^{(1-\beta)/\beta}$. If (i) were correct, Eq. (6.12) would imply $p(\delta) \approx p(\frac{1}{2})(\frac{\delta}{2\beta})^{(1-\beta)/\beta}/2\beta^2$ in contradiction to (i). Assuming (ii), Eq. (6.12) simplifies to $p(\delta) \sim 1/\delta$, which is consistent with (ii) since $(1-\beta)/\beta > -1$. Assuming (iii), Eq. (6.12) leads to

$$p(\delta) \approx \frac{p(\frac{1}{2})}{2\beta^2(1-(2\beta)^{-1/\beta})}\left(\frac{\delta}{2\beta}\right)^{(1-\beta)/\beta}. \tag{6.13}$$

Numerical experiments Zou (2012) show that the latter is the correct approximation of the invariant density since the sampling distribution approaches this form for large N. In contrast, the other possible form $p(\delta) \sim 1/\delta$ only occurs in intermittent phases after a step in which the orbit was mapped very close to zero.

For $\beta > 1$, the above general results imply that the invariant density $p(x)$ has two peaks at $x = 0$ and $x = 1$ with $p(\delta) \sim p(1-\delta) \sim \delta^{(1-\beta)/\beta}$ for small δ. So the above theory predicts that also the degree distribution $P(k)$ displays a power-law with the exponent

$$\gamma = \frac{\beta}{\beta - 1}, \tag{6.14}$$

which agrees very precisely with numerical results Zou (2012).

For $\beta = 1$ (tent map), the nodes of the RN are instead uniformly distributed, and the degree distribution derived from Eq. (6.2) thus takes a Poissonian form, $P(k) = e^{-\alpha}\alpha^k/k!$ Dall (2002).

Finally, for the critical value $\beta = \frac{1}{2}$ (cusp map), we get $p(x) = 2 - 2x$, which again implies the presence of a power-law in $P(k)$ with the critical exponent $\gamma = 1$.

This example demonstrates that the scaling exponents of the degree distribution of RNs are not directly related with a possible fractal dimension of the system, but can (for one-dimensional discrete systems) be exclusively derived from the general shape of the invariant density. Specifically, in the generalized logistic map, the value of γ explicitly depends on β in general (Eq. 6.14). Moreover, we find that power-laws are absent for the tent map ($\beta = 1$), although this system has the same box-counting and correlation dimension as for $1 < \beta \leq 2$. We will discuss about this point in detail in the next section.

6.3 Power-laws and fixed points in 2D flows

If the invariant density p of a flow $\dot{x} = \Phi(x)$ has only one peak, at a fixed point x_0 with $\Phi(x_0) = 0$ whose Laplacian matrix $L = D\Phi(x_0)$ is non-singular, and L's complex eigenvalues fulfil certain conditions, one can analytically show that the degree distribution will follow a slowly decaying power-law of the form $P(k) \propto 1/k$. It is worth pointing out that we follow the traditional linearization theorem of dynamical systems theory to understand how the linearized dynamics determines the existence of power laws of the resulted density function $p(x)$, hence, degree distribution $P(k)$. For two-dimensional flows, one can distinguish the following cases for the two eigenvalues λ_1, λ_2, leading to different types of local flow and corresponding density close to x_0:

(i) If both λ_i are equal and real, the local flow is radial, so that a small volume element dV is stretched in both radial and tangential directions. Then $p(x) \approx g(\phi)/r^2$ with some function g, where r, ϕ are polar coordinates centered at x_0. Independent from g, Eq. (1) then gives the power-law $P(k) \propto 1/k$.

(ii) Similarly, in the frequent case in which both λ_i are complex with non-zero real part, the local flow is spiralling, and $p(x) \approx g(\phi - a \ln r)/r^2$ with some function g and some $a \neq 0$. Again, independent of g we get a power-law $P(k) \propto 1/k$.

(iii) If both λ_i are purely imaginary, the local flow is approximately circular, so that a small volume element dV is only rotated but not stretched. Hence $p(x) = h(r)$ with some function h, and in this case it will depend on the form of h whether $P(k)$ will follow a power-law. E.g., if $h(r) = r^{-2}$, we get again $P(k) \propto 1/k$.

(iv) If both λ_i are real and $\lambda_1 > 0 > \lambda_2$, the local flow has a hyperbolic shape of the form $y(t) \approx ae^{\lambda_1 t}$ and $z(t) \approx be^{\lambda_2 t}$, where y, z are Cartesian coordinates centered at x_0 with axes parallel to the eigenvectors of L. The local density p is then approximately determined by its values along the two diagonals $y = z$ and $y = -z$ and at four points $(y, z) = (\pm a, 0)$ and $(y, z) = (0, \pm a)$ with some small $a > 0$, but those values will depend on the dynamics farther away from x_0. Again, $P(k)$ might or might not follow a power-law in this case.

(v) Finally, if both λ_i are real and $\lambda_1 > \lambda_2 > 0$ or $0 > \lambda_1 > \lambda_2$, the local flow is also $y(t) \approx ae^{\lambda_1 t}$ and $z(t) \approx be^{\lambda_2 t}$, but this now has a parabolic shape, and p is approximately determined by its values along a small circle centered at x_0. As before, these values depend on the dynamics farther away from x_0, so that $P(k)$ might or might not follow a power-law.

For higher-dimensional flows, eigenvalues are either real or come in complex conjugate pairs, so that a similar case distinction can be made with cases that are basically combinations of the above cases. In particular, for the generic case in which at most one eigenvalue is real, the local flow will be a

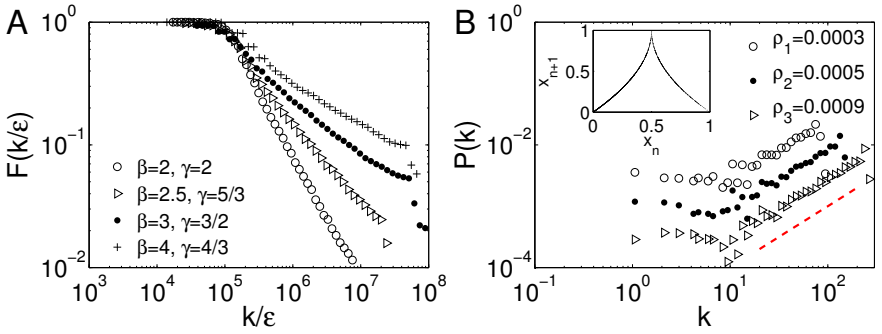


Fig. 6.3: (Colour online) (A): Cumulative degree distribution $F(k)$ of the RNs for the map (6.9) with $\beta = 2, 2.5, 3, 4$. Note that we divide the degree k by ε so that the x-axis is ε-dimensionless. (B) $P(k)$ of the cusp map $\beta = 0.5$. The inset shows the corresponding iterative function $f(x)$. The dashed line of (B) has a slope of 1. The legend indicates the respective link densities ρ.

superposition of spirals and at most one radial component, which will still lead to $P(k) \propto 1/k$.

6.4 Power-law scaling versus fractal dimension

To study the relationship between the scaling exponent and local dimension, let us first focus on discrete-time systems and consider the generalised logistic map (Eq. 6.9) Lyra (1998) again. For general $\beta > 0$, the unit interval $[0, 1]$ is mapped onto itself by a symmetric function with a maximum of 1 at $x = \frac{1}{2}$, thus having two pre-images for each $x < 1$. For $\beta > 1$, the associated invariant density $p(x)$ has two peaks at $x = 0$ and $x = 1$ with $p(\delta) = p(1-\delta) \sim \delta^{(1-\beta)/\beta}$ for small δ. Hence, the degree distribution $P(k)$ shows a power-law with the exponent $\gamma = \beta/(\beta - 1)$ (Eq. 6.14).

Numerical results shown in Fig. 6.3A for several different values of β agree precisely with Eq. (6.14). In contrast, for $\beta = 1$ the nodes are uniformly distributed, and the degree distribution derived from Eq. (6.2) is Poissonian, $P(k) = e^{-\alpha}\alpha^k/k!$ Dall (2002). For $\beta = \frac{1}{2}$, we get $p(x) = 2 - 2x$, which leads to a specific type of "power-law" in $P(k)$ with $\gamma = -1$ as shown in Fig. 6.3B. These results imply that the scaling exponent is *not* simply related to the fractal dimension: the attractor has the box-counting dimension $D_0 = 1$ independently of β, whereas γ changes with varying β (Eq. (6.14)). However, the correlation dimension D_2 also depends on β ($D_2 = 1$ for $\beta \leq 2$, and $D_2 = 2/\beta$ if $\beta > 2$ Sprott (2003)), *i.e.,* there is an indirect relationship between γ and D_2 for certain special cases. The different behaviour of the mentioned dimensions results from the fact that D_0 exclusively considers the

number of boxes required for covering the attractor, but not their individual probability masses as D_2 and other notions of fractal dimensions do.

Turning to *continuous-time systems,* we next compare the above findings with those for some *discretised* standard examples. On the one hand, for the Rössler system, we consider the successive x-values when passing the Poincaré section at $y = 0$ with $\dot{y} < 0$. As shown in the inset of Fig. 6.4A, the resulting first return map has a shape similar to the case of $\beta = 1.87$ in Eq. (6.9). Hence, we expect a power-law with the exponent $\gamma \approx 1.87/(1.87 - 1) = 2.15$. The invariant density has several dominant peaks, which are together responsible for the power-law observed in Fig. 6.2A with γ indeed close to 2.15. In fact, $P(k)$ is a mixture of individual power-laws corresponding to the individual peaks of $p(x)$, whose exponents are all roughly the same. On the other hand, for the Lorenz system, we obtain a one-dimensional map by studying the local maxima z^n_{max} of z for successive cycles Lorenz (1963), *i.e.,* mapping z^n_{max} to z^{n+1}_{max} (inset of Fig. 6.4B). For $r = 28$, this first return map has a similar shape as Eq. (6.9) for $\beta = 0.5$ (inset of Fig. 6.3B), but the corresponding density is bell-shaped without a peak. Indeed, we do not observe a power-law for $P(k)$ in this case (Fig. 6.2B). However, increasing r changes the shape of $p(x)$. For example, at $r = 90$ (Fig. 6.4C) the density has peaks at several points, explaining the observed power-law in Fig. 6.2C. These results demonstrate that scaling is only present for a certain range of r values. Given the rich bifurcation scenario of the Lorenz system a corresponding detailed study is beyond the scope of this work, but will be a subject of future research. A similar behaviour as for the discretised Lorenz system can be observed for the Hénon map, though the marginal invariant density of the x-component has a more complex structure (Fig. 6.4D).

While there is no unambiguous relationship between γ and the fractal dimension already for discrete systems, the situation becomes even more complicated for *continuous-time* systems which are not discretised via a Poincaré section or otherwise. For two-dimensional flows $\dot{x} = \Phi(x)$ with only one peak in $p(x)$, the respective type of behaviour depends on the eigenvalues of the Jacobian $D\Phi(x_0)$ at the fixed point x_0 as well as on the shape of $p(x)$. Specifically, in many cases (that have been further discussed in sec. 6.3) the existence of a power-law for $P(k)$ cannot be evaluated easily, whereas in other cases, one can analytically derive a power-law with a very small exponent of $\gamma = 1$. In turn, the following numerical results suggest that there are also examples displaying a distinct relationship between γ and some suitably defined local dimension:

For the Rössler system in the regime of screw-type chaos with a homoclinic point at the origin fulfilling the Shilnikov condition Shilnikov (1970), Gaspard (1983), the invariant density is dominated by its peak at the origin. The degree distribution $P(k)$ of the corresponding RN shows a power-law with $\gamma \approx 1.33$, which agrees fairly well with the ϑ-capacity dimension D^{ϑ}_0 defined in Farmer (1983) (Fig. 6.5). We also observe similar scaling laws for both numerical model and experimental data (output intensities) of a single-mode CO_2

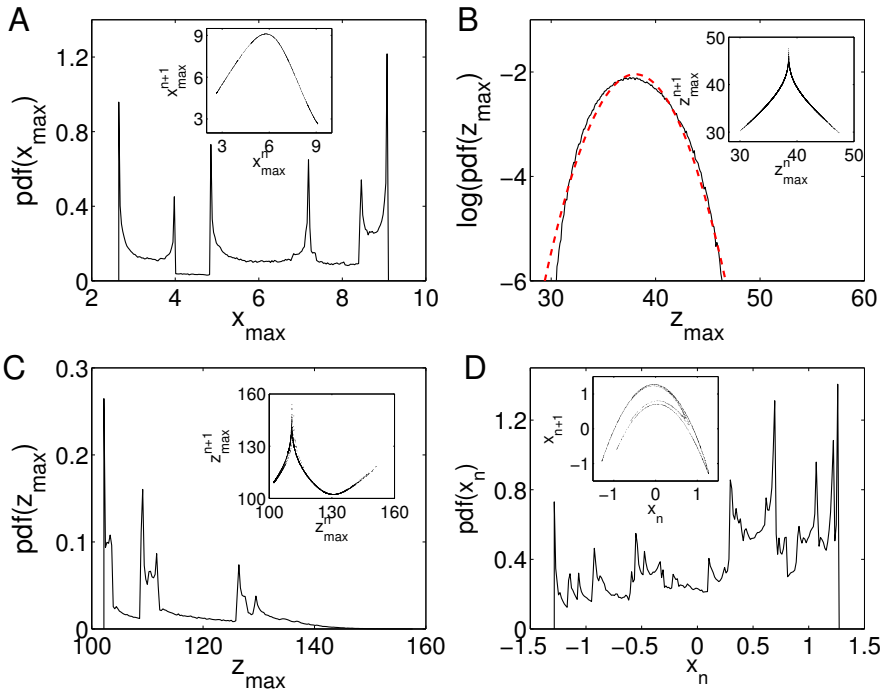


Fig. 6.4: (Colour online) Probability density function (PDF) of (A) the successive x values passing the plane of $y = 0$ of the Rössler attractor, (B,C) PDF of z_{max} for the Lorenz system of $r = 28$ (B) and $r = 90$ (C), and (D) marginal density of the x component of the Hénon map. The insets show the associated first return plot. The dashed curve in (B) shows a fit by a Gaussian distribution.

laser Pisarchik (2001). The underlying system has a saddle-focus S embedded in the chaotic attractor (Fig. 6.6A) which causes a spiking dynamics Channell (2007), Arecchi (1987). The attractor is dominated by a homoclinic orbit emerging from and converging to S. The degree distributions $P(k)$ resulting from both model and experimental data suggest power-laws with $\gamma \approx 1.35$ (Fig. 6.6B), which qualitatively agrees well with the point-wise dimension of the attractor around S. Finally, similar results can be obtained for a predator-prey food-chain model with four competing species Vano (2006), which also displays homoclinic chaos, where we observe $\gamma \approx 1.9$ in agreement with D_0^{ϑ} (Fig. 6.7). This variety of examples underlines the general importance and wide applicability of our findings.

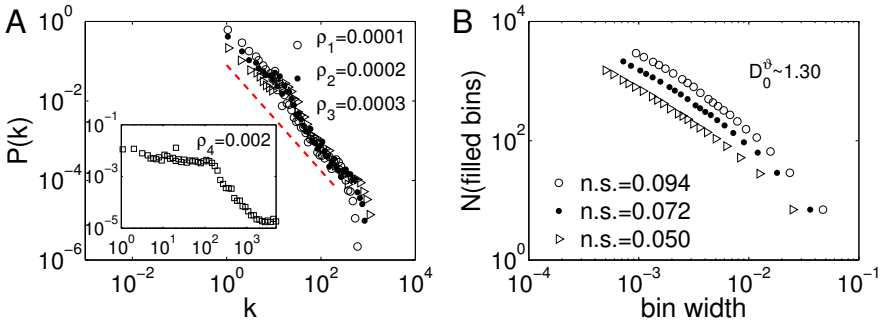


Fig. 6.5: (Colour online) (A): $P(k)$ for the RN of the Rössler system with screw-type chaos Gaspard (1983). Three different link densities are chosen for illustration as indicated by the legend, a larger link density ρ_4 does not produce a power-law (inset). The slope of the dashed line is -1.33. (B): ϑ-capacity dimension $D_0^\vartheta \approx 1.30$ in three small cuboidal neighbourhoods of different size (in terms of phase space distance in each coordinate direction, see the legend).

Fig. 6.6: (Colour online) Experimental laser data: (A) Phase portrait in subspace of (x_1, x_2, x_6), where the saddle focus S is at the most dense region. (B) $P(k)$ of RNs from model data for three different link densities (see the legend); a larger link density ρ_4 does not lead to a power-law (inset). (C) $P(k)$ of RNs from experimental data (inset: cumulative distribution $F(k)$). (D) Point-wise dimension $D_1^S = 1.35$. All dashed lines have slope -1.35.

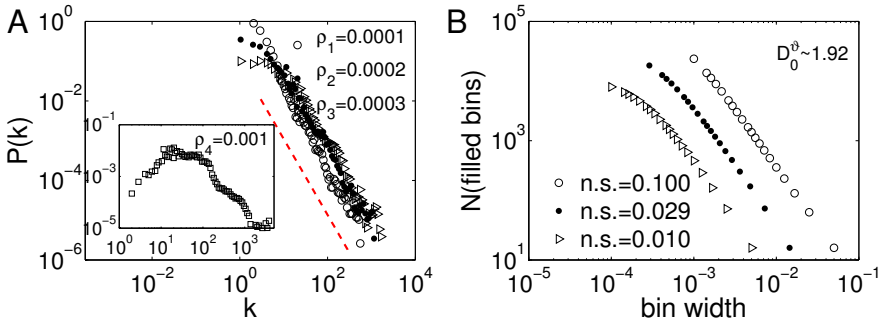


Fig. 6.7: (Colour online) As in Fig. 6.5 for the Lotka-Volterra system Vano (2006) with homoclinic chaos. The slope of the dashed line in panel (A) is -1.90.

6.5 Technical aspects

In general, we have to make two cautionary notes on the numerical study of scaling laws in RNs.

First, in the previous continuous-time examples, the presence of power-laws with the numerically estimated exponents (see above) cannot be rejected on a 90% significance level using Kolmogorov-Smirnov tests (as shown in Fig. 6.8 in sec. 6.5.1). However, the alternative of a power-law with $\gamma = 1$ can also not be rejected at the same level. Hence, power-laws with qualitatively different exponents describe the data comparably well. It remains an open problem to determine the correct γ. We note that this is a general problem when evaluating hypothetical power-laws from finite data Clauset (2009), Stumpf (2012).

Second, experimental data often consist of only one measured variable. Hence, a reconstruction of the associated phase space trajectory is necessary prior to RN analysis, *e.g.,* by time-delay embedding Packard (1980). Like estimates of dynamical invariants or complexity measures Letellier (2005, 2006), the power-law behaviour of $P(k)$ can depend on the particular observable, because different coordinates of a dynamical system often have different marginal densities. Specifically, embedding theorems ensure *topological* invariance (*i.e.,* properties of the dynamical system that do not change under smooth coordinate transformations are preserved), but no *metric* invariance of the attractor's geometry including $p(x)$. For example, the logistic map and the tent map ($\beta = 2, 1$ in Eq. (6.9), respectively) are topologically equivalent under the transformation $x \mapsto \sin^2(\pi x/2)$, but the different invariant densities with respect to their original coordinates (that have been used for constructing the RNs from metric distances in their respective phase spaces) lead to distinct scaling exponents γ (Fig. 6.3A). In other words, the power-law exponent γ is not invariant with respect to general smooth coordinate transformations that do not leave the geometric shape of the system's invariant density qualitatively the same.

6.5.1 Estimation of scaling exponents

Assume the probability density function of a random variable X is approximately $f(x) \approx Ax^{-\gamma}$ for $x > x_c$, with $\gamma = 1$ and $A > 0$. Here, x_c is a lower cutoff that is related to some typical minimum scale below which the supposed power-law model is not meaningful anymore Clauset (2009). Then the cumulative distribution function (CDF) is undefined since $\int_x^\infty dy/y$ does not converge. In tests using the CDF, one can then try to use instead the function $F_m(x) = \int_x^m dy\, A/y = A(\ln m - \ln x)$ with some value m at least as large as all observations, e.g., putting $m = \max_i x_i$. The empirical counterpart of $F(x)$ is then $\hat{F}(x) = |\{i : x_i > x\}|/n_c$ with $n_c = |\{i : x_i > x_c\}|$.

Hence, a simple Kolmogorov-Smirnov-like test uses the statistic $D_s = \sup_{x\in[x_c,m]} |\hat{F}(x) - F(x)|$ which approximately follows a Kolmogorov distribution with n_c degrees of freedom if the data come from the power-law, so that the test is whether $\sqrt{n_c} D_s$ exceeds the critical value K_α of the Kolmogorov

distribution corresponding to the sought significance level α. As suggested in Clauset (2009), x_c could be determined from the data by choosing the value that minimizes D_s. In the modified test described above, however, one also needs to estimate the constant A, and both can be done simultaneously by finding the combination of x_c and A that minimizes D_s.

If $\gamma > 1$ but still very close to 1, the CDF exists but it might still be more robust to use the same method as above, now using $F(x) = \int_x^m dy\, Ay^{-\gamma} = A(x^{1-\gamma} - m^{1-\gamma})/(\gamma - 1)$.

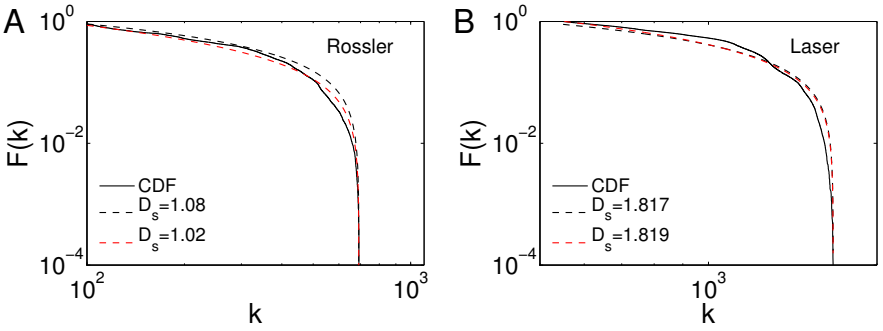


Fig. 6.8: (Color online) Cumulative degree distribution $F(k)$ for (A) Rössler system of screw chaos and (B) CO_2 laser model. Fitting by power law of exponents of both 1 (black dashed line) and local dimension value (red dashed line) show the same discriminative strength as indicated by Kolmogorov-Smirnov test D_s.

6.5.2 Selection of dynamical variable

Data obtained by experiment often consists of only one scalar variable, hence a reconstruction of its associated phase space trajectory is required before performing RN analysis, e. g., by time-delay embedding Packard (1980). The power-law behavior of $P(k)$ is influenced by which observable is chosen for embedding. To illustrate this, we consider a periodic solution in a two-dimensional phase space, which is obtained by integrating the van der Pol oscillator $\ddot{x} - \mu(1 - x^2)\dot{x} + x = 0$ for the nonlinear parameter $\mu = 10$ (we also qualitatively compare the difference to the case of $\mu = 1$ in the insets of Fig. 6.9). The two components of x and y are functionally independent, having different marginal PDF in the corresponding subspace as shown in Fig. 6.9(C, D), respectively. This results in quite different contributions to the degree distribution $P(k)$ in joint space. The final form of $P(k)$ is also influenced by the metric which we used to define the phase space distance (e.g., in Fig. 6.1). More specifically, $P(k)$ consists of two well separated branches if

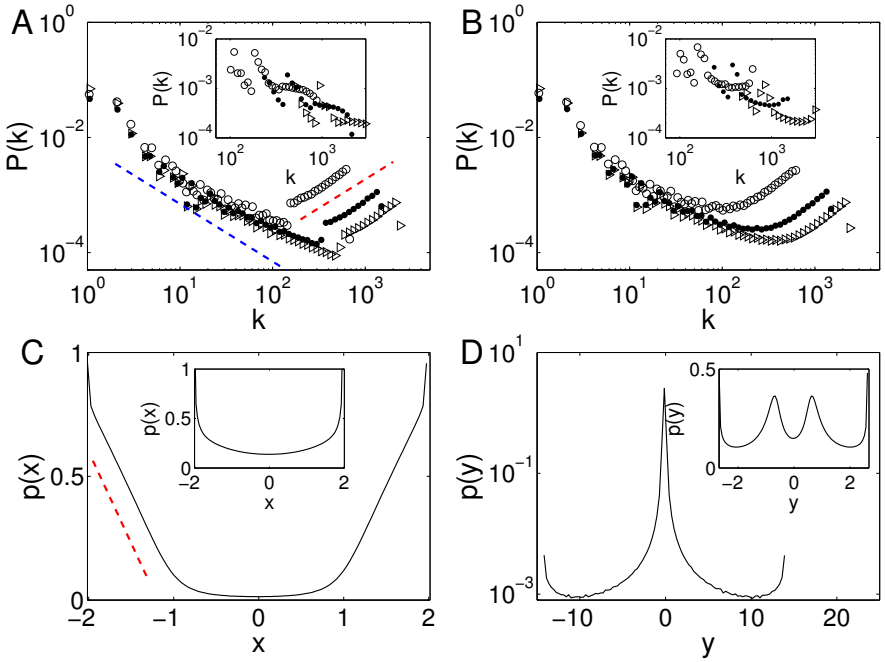


Fig. 6.9: (Color online) $P(k)$ for the van der Pol oscillator for $\mu = 10$ with (A) maximum norm and (B) Euclidean distance. Three curves of link densities are chosen for illustration, $\rho_1 = 0.2\%(\circ)$, $\rho_2 = 0.5\%(\bullet)$ and $\rho_3 = 0.7\%(\triangleright)$. In addition, the marginal PDFs (C) of the x component and (D) of the y component are shown, respectively. The insets show the case of $\mu = 1$. The dashed line of (C) has a slope of absolute value of $\frac{3}{4}$.

the maximum norm is applied (Fig. 6.9A). The decreasing trend of the first branch (characterized by $\gamma \approx 2$) mainly results from the fast decay of the marginal PDF of the y component, while the positive trend of the second branch with a slope of 1 comes from the almost linear regime of the marginal PDF of x. It is not too difficult to understand the mixture property if the Euclidean metric is applied. The two branches merge and the power-laws are smeared out (Fig. 6.9B). Since the system shows strong relaxation oscillations for $\mu = 10$, the degree distribution $P(k)$ of both x and y in the joint space can be well approximated by a combination of a linear model (for instance, $\beta = 1/2$ in Eq. (6.9)) and an exponential model with some appropriate coordinate transformations.

6.6 Conclusions

In summary, we have reported an interesting novel aspect of the geometrical organisation underlying the dynamics of many complex systems in physics and beyond. Specifically, we have provided an analytical explanation of the emergence of power-laws in recurrence networks constructed from sampled time series based on the theory of random geometric graphs. Unlike for comparable complex network approaches Zhang (2006), Lacasa (2008), this scaling is not simply related to the system's fractal dimension, but requires the presence of power-law shaped singularities of the invariant density and a feasible choice of the considered spatial scale ε. We emphasise that dimensions are defined in the limit of $\varepsilon \to 0$ and practically estimated by a series of ε values, whereas the power-law exponent γ of the RN appears for each sufficiently small ε *individually*. Note that in contrast to the degree, the transitivity properties of RNs have a direct relationship with attractor dimension Donner (2011).

In comparison with the invariant density itself, fractal dimensions are a rather specific characteristic. In particular, they do not simply describe the whole system (as the invariant density itself does), but quantify density variations on the attractor viewed at different spatial scales ε Farmer (1983), Halsey (1986). Conversely, the scaling exponent γ directly characterises a power-law decay of the density in phase space independent of a specific scale. In this spirit, both fractal dimension and scaling exponent γ capture conceptually different aspects of the geometric organisation of a dynamical system in its phase space. However, although there is no general relationship between γ and fractality, in some special cases the power-law exponent coincides with some notion of dimension. This has been demonstrated for several example systems as well as experimental data. In turn, we have found that in other cases the value of γ drops to 1. There is a need for further studies involving both additional model systems and experimental examples beyond those discussed in this work in order to better understand this complex relationship between power-law degree distributions and fractal scaling (i.e. under which general conditions related to a system's structural organisation both scaling exponent and fractal dimension coincide), particularly in continuous dynamical systems.

From a conceptual perspective, we would like to remark that studying a single scalar property like the scaling exponent of a recurrence network or the fractal dimension cannot provide a complete view on the structural organisation of a nonlinear complex system. Specifically, both characteristics capture distinct and complementary features related to the probability density of the invariant measure. In this spirit, the power-law exponent γ quantifies a fundamental property that has not been explicitly studied so far. Because its relationship with the features of possible singularities of the invariant density is intuitive (i.e. the emergence of power-law degree distributions has some clear physical meaning), one particular strength of studying the degree dis-

tribution of recurrence networks is that it potentially allows identifying the presence of such singularities in complex situations (*e.g.,* for observational data).

Power law distributions of various physical parameters in nature are hallmarks of many systems showing self-organized criticality (SOC). The most widely known example is the distribution of earthquake magnitudes which has the so-called Gutenberg-Richter law Gutenberg (1944). The most recent application of the concepts of SOC was initially to disclose sandpile avalanches at a critical angle of response Bak (1987, 1988), and has been enthusiastically applied across fields as diverse as geophysics Bak (1989), evolutionary biology Sneppen (1995), Killingback (1998), economics Mantegna (1995), Farmer (2004), solar physics Aschwanden (2011), and neurosciences Poil (2008) and many others. The relationships between our findings of power laws of the degree distribution $P(k)$ of the associated recurrence networks and the SOC in general are a topic of our ongoing work. More specifically, we intend to performing recurrence analysis for a slowly-driven SOC system which has been demonstrated to have avalanches in the occurrence frequecy distributions and parameter correlations Aschwanden (2012). We expect that our recurrence network approaches could shed some light on the still ambiguous understanding of the maximal complexity while the parameter of the system is tuned on the edge of chaos Langton (1990), Mitchell (1993, 1994).

Acknowledgements

This work has been partially funded by the Leibniz society (project ECONS) and the Federal Ministry for Education and Research (BMBF) via the Potsdam Research Cluster for Georisk Analysis, Environmental Change and Sustainability (PROGRESS), the National Natural Science Foundation of China (Grant No. 11135001), and the Hong Kong Polytechnic University Postdoctoral Fellowship. We acknowledge the discussion with J. D. Farmer, R. V. Donner, J. F. Donges, N. Marwan, and M. Small. We also thank R. Meucci and S. Euzzor for providing the experimental data.

References

Albert, R., Barabasi, A.L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(1), 47–97 (2002)

Arecchi, F.T., Meucci, R., Gadomski, W.: Laser dynamics with competing instabilities. Phys. Rev. Lett. **58**, 2205 (1987)

Arenas, A., Diaz-Guilera, A., Kurths, J., Moreno, Y., Zhou, C.: Synchronization in complex networks. Phys. Rep. **469**(3), 93–153 (2008)

Aschwanden, M.: Self-Organized Criticality in Astrophysics, 1st edition edn. Springer Praxis Books (2011)
Aschwanden, M. J.: A statistical fractal-diffusive avalanche model of a slowly-driven self-organized criticality system. Astronomy and Astrophysics **539**, A2 (2012)
Bak, P., Tang, C.: Earthquakes as a self-organized critical phenomena. J. Geophys. Res. **94**, 15,635–15,637 (1989)
Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality: An explanation of the $1/f$ noise. Phys. Rev. Lett. **59**, 381–384 (1987)
Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality. Phys. Rev. A **38**, 364–374 (1988)
Barthélemy, M.: Spatial networks. Phys. Rep. **499**, 1 – 101 (2011)
Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U.: Complex networks: Structure and dynamics. Phys. Rep. **424**(4-5), 175–308 (2006)
Channell, P., Cymbalyuk, G., Shilnikov, A.: Origin of bursting through homoclinic spike adding in a neuron model. Phys. Rev. Lett. **98**, 134,101 (2007)
Clauset, A., Shalizi, C.R., Newman, M.E.J.: Power-law distributions in empirical data. SIAM Rev. **51**, 661 (2009)
Dall, J., Christensen, M.: Random geometric graphs. Phys. Rev. E **66**(1), 016,121 (2002)
Donges, J.F., Donner, R.V., Trauth, M.H., Marwan, N., Schellnhuber, H.J., Kurths, J.: Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. Proc. Natl. Acad. Sci. **108**, 20,422–20,427 (2011)
Donges, J.F., Heitzig, J., Donner, R.V., Kurths, J.: Analytical framework for recurrence network analysis of time series. Phys. Rev. E **85**, 046,105 (2012)
Donges, J.F., Schultz, H.C.H., Marwan, N., Zou, Y., Kurths, J.: Investigating the topology of interacting networks. Eur. Phys. J. B **84**, 635–651 (2011)
Donner, R.V., Heitzig, J., Donges, J.F., Zou, Y., Marwan, N., Kurths, J.: The geometry of chaotic dynamics - a complex network perspective. Eur. Phys. J. B **84**(4), 653–672 (2011)
Donner, R.V., Zou, Y., Donges, J.F., Marwan, N., Kurths, J.: Ambiguities in recurrence-based complex network representations of time series. Phys. Rev. E **81**(1), 015,101(R) (2010)
Donner, R.V., Zou, Y., Donges, J.F., Marwan, N., Kurths, J.: Recurrence networks – A novel paradigm for nonlinear time series analysis. New J. Phys. **12**(3), 033,025 (2010)
Eckmann, J.P., Ruelle, D.: Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. **57**, 617 (1985)
Farmer, J.D., Lillo, F.: On the origin of power law tails in price fluctuations. Quant. Finance **4**, 7–11 (2004)
Farmer, J.D., Ott, E., Yorke, J.A.: The dimension of chaotic attractors. Physica D **7**(1-3), 153–180 (1983)
Gao, J., Buldyrev, S., Stanley, H., Havlin, S.: Networks formed from interdependent networks. Nature Physics **8**, 40–48 (2012)
Gaspard, P., Nicolis, G.: What can we learn from homoclinic orbits in chaotic dynamics? J. Stat. Phys. **31**, 499 (1983)
Grassberger, P., Procaccia, I.: Characterization of strange attractors. Phys. Rev. Lett. **50**(5), 346–349 (1983)
Gutenberg, B., Richter, C.F.: Frequency of earthquakes in california. Bull. Seism. Soc. Am. **34**, 185–188 (1944)
Halsey, T.C., Jensen, M.H., Kadanoff, L.P., Procaccia, I., Shraiman, B.I.: Fractal measures and their singularities: The characterization of strange sets. Phys. Rev. A **33**, 1141 (1986)
Herrmann, C., Barthélemy, M., Provero, P.: Connectivity distribution of spatial networks. Phys. Rev. E **68**(2), 026,128 (2003)

Kantz, H., Schreiber, T.: Nonlinear time series analysis. Cambridge University Press (1997)
Killingback, T., Doebeli, M.: Self-organized criticality in spatial evolutionary game theory. Journal of Theoretical Biology **191**, 335–340 (1998)
Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuno, J.C.: From time series to complex networks: The visibility graph. Proc. Natl. Acad. Sci. **105**(13), 4972–4975 (2008)
Langton, C.G.: Computation at the edge of chaos: Phase transitions and emergent computation. Physica D Nonlinear Phenomena **42**, 12–37 (1990)
Letellier, C.: Estimating the shannon entropy: Recurrence plots versus symbolic dynamics. Phys. Rev. Lett. **96**(25), 254,102 (2006)
Letellier, C., Aguirre, L.A., Maquet, J.: Relation between observability and differential embeddings for nonlinear dynamics. Phys. Rev. E **71**, 066,213 (2005)
Lorenz, E.: Deterministic nonperiodic flow. J. Atmos. Sci. **20**, 130 (1963)
Lyra, M.L., Tsallis, C.: Nonextensivity and multifractality in low-dimensional dissipative systems. Phys. Rev. Lett. **80**, 53 (1998)
Mantegna, R.N., Stanley, H.E.: Scaling behaviour in the dynamics of an economic index. Nature **376**, 46–49 (1995)
Marquet, P.A., Quiñones, R.A., Abades, S., Labra, F., Tognelli, M., Arim, M., Rivadeneira, M.: Scaling and power-laws in ecological systems. The Journal of Experimental Biology **208**, 1749–1769 (2005)
Marwan, N., Donges, J.F., Zou, Y., Donner, R.V., Kurths, J.: Complex network approach for recurrence analysis of time series. Phys. Lett. A **373**(46), 4246–4254 (2009)
Marwan, N., Romano, M.C., Thiel, M., Kurths, J.: Recurrence Plots for the Analysis of Complex Systems. Phys. Rep. **438**(5–6), 237–329 (2007)
Mitchell, M., Crutchfield, J.P., Hraber, P.T., Cowan, I.G., Pines, D., (editors, D.M., Metaphors, C.: Dynamics, computation, and the "edge of chaos": A re-examination. In: Complexity:Metaphors, Models, and Reality, pp. 497–513. Addison-Wesley (1994)
Mitchell, M., Hraber, P.T., Crutchfield, J.P.: Revisiting the edge of chaos: Evolving cellular automata to perform computations. Complex Systems **7**, 89–130 (1993)
Newman, M.E.J.: The structure and function of complex networks. SIAM Rev. **45**(2), 167–256 (2003)
Packard, N.H., Crutchfield, J.P., Farmer, J.D., Shaw, R.S.: Geometry from a Time Series. Phys. Rev. Lett. **45**(9), 712–716 (1980)
Pisarchik, A.N., Meucci, R., Arecchi, F.T.: Theoretical and experimental study of discrete behavior of shilnikov chaosin a co2 laser. Eur. Phys. J. D **13**, 385 (2001)
Poil, S.S., van Ooyen, A., Linkenkaer-Hansen, K.: Avalanche dynamics of human brain oscillations: Relation to critical branching processes and temporal correlations. Human Brain Mapping **29**(7), 770–777 (2008)
Poincaré, H.: Sur la problème des trois corps et les équations de la dynamique. Acta Mathematica **13**(1), A3–A270 (1890)
Shilnikov, L.: A contribution to the problem of the structure of an extended neighborhood of a rough equilibrium state of saddle-focus type. Math. USSR Sbornik **10**, 91 (1970)
Sneppen, K., Bak, P., Flyvbjerg, H., Jensen, M.H.: Evolution as a self-organized critical phenomenon. Proc. Natl. Acad. Sci. **92**(11), 5209–5213 (1995)
Sprott, J.: Chaos and Time Series Analysis. Oxford University Press (2003)
Stumpf, M.P.H., Porter, M.A.: Critical truths about power laws. Science **335**(6069), 665–666 (2012)
Vano, J., Wildenberg, J., Anderson, M., Noel, J., Sprott, J.: Chaos in low-dimensional lotka–volterra models of competition. Nonlinearity **19**, 2391 (2006)

Xu, X., Zhang, J., Small, M.: Superfamily phenomena and motifs of networks induced from time series. Proc. Natl. Acad. Sci. **105**(50), 19,601–19,605 (2008)
Zhang, J., Small, M.: Complex network from pseudoperiodic time series: Topology versus dynamics. Phys. Rev. Lett. **96**(23), 238,701 (2006)
Zou, Y., Heitzig, J., Donner, R.V., Donges, J.F., Farmer, J.D., Meucci, R., Euzzor, S., Marwan, N., Kurths, J.: Power-laws in recurrence networks from dynamical systems. EPL (Europhysics Letters) **98**(4), 48,001 (2012)

Chapter 7
SOC computer simulations

Gunnar Pruessner

Abstract The following chapter provides an overview of the techniques used to understand Self-Organised Criticality (SOC) by performing computer simulations. Those are of particular significance in SOC, given its very paradigm, the BTW (Bak-Tang-Wiesenfeld) sandpile, was introduced on the basis of a process that is conveniently implemented as a computer program. The chapter is divided into three sections: In the first section a number of key concepts are introduced, followed by four brief presentations of SOC models which are most commonly investigated or which have played an important part in the development of the field as a whole. The second section is concerned with the basics of scaling with particular emphasis of its rôle in numerical models of SOC, introducing a number of basic tools for data analysis such as binning, moment analysis and error estimation. The third section is devoted to numerical methods and algorithms as applied to SOC models, addressing typical computational questions with the particular application of SOC in mind. The present chapter is rather technical, but hands-on at the same time, providing practical advice and even code snippets (in C) wherever possible.

7.1 Introduction

The concept of Self-Organised Criticality (SOC)[1] was introduced by Bak *et al.* (1987) on the basis of a computer model, the famous BTW Sandpile. The notion of "computer model" and "simulation" used here is subtle and can be misleading. Often the models are not meant to mimic a particular (natural) phenomenon, but are intended to capture merely what is considered

Imperial College London, Department of Mathematics, e-mail: g.pruessner@imperial.ac.uk

[1] A more extensive review on the present subject area can be found in (Pruessner 2012b).

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

to be the *essential* interaction observed in a natural phenomenon. Per Bak in particular, had the tendency to name models according to their appearance rather than their purpose and so the "Sandpile Model" may not have been envisaged to display the dynamics of a sandpile. The situation is clearer in the case of the "Forest Fire Model" (Bak *et al.* 1990), which was developed as a model of turbulence much more than as a model of fires in woods.

In particular in the early days of SOC modelling, the models were sometimes referred to as "cellular automata" Olami *et al.* (1992), Lebowitz *et al.* (1990), which caused some consternation (*e.g.* Grassberger 1994), as cellular automata normally have discrete states and evolve in discrete time steps according to deterministic rules in discrete space (*i.e.* a lattice). The term "coupled map lattice" (Kaneko 1989) can be more appropriate for some models, such as the Olami-Feder-Christensen Model descusssed below (discrete space, continuous state and possibly continuous time).

The terminology of "numerical modelling" has always been somewhat confusing. Many of the models considered in SOC do not model a natural phenomenon and so their numerical implementation is not a "numerical simulation" in the sense that they mimic the behaviour of something else. There are notable exceptions, however, such as the Forest Fire Model (Bak *et al.* 1990) mentioned above and the Oslo ricepile model (Christensen *et al.* 1996). SOC models generally are not "models of SOC", rather they are algorithmic prescriptions or "recipes" for a (stochastic) process that is believed to exhibit some of the features normally observed in other SOC models. In that sense, the terminology of terms like "SOC models" and "simulation" or even "simulating an SOC model" is misleading — most of these models are not simplified versions or idealisations of some physical process or anything else that is readily identified as "SOC", but recipes to produce some of the behaviour expected in an SOC system.

To this day, a large fraction of the SOC community dedicate their research to computer models. Initially, the motivation (*e.g.* Zhang 1989, Manna 1991) was to find models displaying the same universal behaviour as the BTW (Bak-Tang-Wiesenfeld) Sandpile. This was followed by an era of proliferation, when many new models, belonging to new universality classes where developed. More recently, in a more reductionistic spirit, new models are mostly developed to isolate the rôle of particular features and to extract and identify their effect (*e.g.* Tadić and Dhar 1997). A lot of numerical research into SOC nowadays happens "en passant", as SOC is identified in a model for a phenomenon that originally was not considered to be related to SOC (*e.g.* Burridge and Knopoff 1967).

Virtually all SOC (computer) models consist of degrees of freedom interacting with (nearest) neighbours located on a lattice. The degrees of freedom may be parameterised by continuous or discrete variables, in the following denoted $z_{\mathbf{n}}$, where $\mathbf{n}$ is a position vector on the lattice. A **slow, external driving mechanism** (in short, **external drive**) slowly loads the system, *i.e.* the local variables are slowly increased, also referred to as "charging a

site". That might happen uniformly (sometimes called **global drive**) or at individual lattice sites (sometimes called **point drive**). The driving might happen at randomly chosen points or by random increments, both of which is in the literature referred to as **random driving**. The dynamics of an SOC model is **non-linear**, *i.e.* there is no linear equation of motion that would describe their dynamics.[2] The response of the system is triggered by a local degree of freedom overcoming a **threshold**, beyond which **relaxation** and thus interaction with other degrees of freedom and the outside world takes place. A site where that happens is said to **topple** and to be **active**. The interaction might lead to one of the neighbours exceeding its threshold in turn, triggering another relaxation event. The totality of the relaxations constitutes an **avalanche**. When the avalanche has finished, *i.e.* there are no active sites left, the system is in a state of **quiescence**. In SOC models, driving takes place only in the quiescent state (separation of time scales, below). If the external drive acts at times when an avalanche is running, it might lead to a continuously running avalanche (e.g. Corral and Paczuski 1999).

In many models the degree of freedom at every site measures a resource that is **conserved** under the dynamics. To balance the external drive, in most models **dissipation** has to take place in some form: **Bulk dissipation** takes place when the resource can get lost in the local interaction. **Boundary dissipation** refers to the situation when the resource is lost only in case a boundary site relaxes. The necessary flux of the resource towards the boundaries has been suggested as some of the key mechanisms in SOC (Paczuski and Bassler 2000b). In some models, such as the Bak-Sneppen Model (Bak and Sneppen 1993) or the Forest-Fire-Models (Henley 1989, Bak *et al.* 1990, Drossel and Schwabl 1992a), no (limited) resource can be identified and therefore the notion of dissipation and conservation is not meaningful.

The question whether conservation is a necessary ingredient of SOC has driven the evolution of SOC models in particular during the 1990s. In fact, early theoretical results by Hwa and Kardar (1989a) suggested that bulk dissipation would spoil the SOC state. Models like the OFC Model (Olami *et al.* 1992, also Bak and Sneppen 1993, Drossel and Schwabl 1992a) questioned that finding. Different theoretical views have emerged over time: Lauritsen *et al.*'s (1996) self-organised branching process (Zapperi *et al.* 1995) contains dissipation as a *relevant* parameter which has a limitting effect on the scaling behaviour. Juanico *et al.* (2007) restored the SOC state of the self-organised branching process by implementing a mechanism that compensates for the non-conservation by a "matching condition" not dissimilar from the mechanism used in the mean-field theory by Pruessner and Jensen (2002b). That, in turn, was labelled by Bonachela and Muñoz (2009) as a form of tuning.

[2] It is very instructive to ask why a non-linearity is such a crucial ingredient. Firstly, if all interactions were linear, one would expect the resulting behaviour to correspond to that of a solvable, "trivial" system. Secondly, linearity suggests additivity of external drive and response, so responses would be expected to be proportional to the drive, a rather boring behaviour, not expected to result in scale invariance.

More recent field-theoretic work (Pruessner 2012b) points at conservation as a symmetry responsible for the cancellation of mass-generating diagrams, an effect that may equally be achieved by other symmetries.

The external drive, the ensuing sequence of avalanches and the evolution of the model from one quiescent state to the next happen on the **macroscopic time scale**, where time typically passes by one unit per avalanche. As the system size is increased, avalanches are expected to take more and more relaxations to complete. Their duration is measured on the **microscopic time scale**. In the thermodynamic limit, *i.e.* at infinite system size, the infinite duration of an avalanche of the microscopic time scale and the finite driving rate on the macroscopic time scale amount to a complete **separation of time scales**. In general, the separation of time scales is achieved in finite systems provided that no driving takes place when any site is active, because the times of quiescence, measured on the microscopic time scale, can be thought of as arbitrarily long. As a result, the avalanching in these systems becomes **intermittent**.

Separation of time scales is widely regarded as *the* crucial ingredient of SOC, maybe because it is conceived (and criticised as such) as a substitute of the tuning found in traditional critical phenomena (also Jensen 1998). In numerical models, it normally enters in a rather innocent way — the system is not driven while an avalanche is running. This, however, requires some global supervision, a "babysitter" (Dickman *et al.* 2000) or a "farmer" (Bröker and Grassberger 1999). In some models the separation of time scales can be implemented explicitly (Bak and Sneppen 1993) in the relaxational rule. What makes the separation of time scales very different from other forms of tuning is that it *eliminates* a dimensionful, finite scale, such as the frequency with which an avalanche is triggered.[3] In traditional critical phenomena, scaling comes about due to the *presence* of a dimensionful, finite energy scale[4], where entropic contributions to the free energy compete with those from the internal energy promoting order. In most SOC models, it is pretty obvious that scaling would break down if time scales were not explicitly separated — avalanches start merging and eventually intermittency is no longer observed (Corral and Paczuski 1999).

SOC models are normally studied at **stationarity**, when all correlations originating from the initial state (often the empty lattice) are negligible. Reaching this point is a process normally referred to as **equilibration**. The equilibration time is normally measured as the number of charges by the external drive required to reach stationarity. For some models, exact upper bounds for the equilibration time are known (Dhar *et al.* 1995, Corral 2004a, Dhar 2004, *e.g.*). In deterministic models, a clear distinction exists between **transient** and **recurrent states**, where the former can appear at most once,

[3] In the field theory of SOC, the cancellation of diagrams occurs precisely when stationarity is imposed for the density of particles resting (and their correlations) in the limit $\omega \to 0$, *i.e.* in the long time limit.

[4] For example $k_B T_c$ in the Ising Model (Stanley 1971).

and the latter with a finite frequency provided the number of states overall is finite. In fact, this frequency is the same for all recurrent states, depending on the driving, which can be at one site only or randomly and independently throughout. A detailed proof of such properties can be cumbersome (Dhar 1999a,b).

The statistics of the avalanches, their size as well as their extent in space and in time, is collected and analysed. SOC is usually said to be found in these models when the statistics displays a **scaling symmetry**, governed by only one upper cutoff which diverges with the system size. In principle, a Gaussian possesses this scaling symmetry,[5] but not a single important SOC model has a Gaussian event size distribution. On the contrary, the avalanche statistics of all models discussed below deviates dramatically from a Gaussian, thus suggesting that avalanches are not the result of essentially independent patches of avalanching sites creating a bigger overall avalanche. Rather, sites are **strongly interacting**, thereby creating the overall event. The purpose of numerical simulations is to characterise and quantify this interaction and its effect, as well as extracting **universal quantities**, which can be compared with those found in other systems.

7.1.1 Observables

As for the methods of analysis, they have matured considerably over the past decades. The initial hunt for $1/f$ noise in temporal signals has given way to the study of event size distributions. As a matter of numerical convenience, these distributions are often characterised using moments, some of which are known exactly. Since the beginning of computational physics, moments and cumulants have been the commonly used method of choice to characterise critical phenomena (Binder and Heermann 1997). It is probably owed to the time of the late 1980's that memory-intensive observables such as entire distributions became computationally affordable and subsequently the centre of attention in SOC.

To this day, the analysis of moments in SOC is still often regarded as an unfortunate necessity to characterise distributions, which are difficult to describe quantitatively. Apart from the historic explanation alluded to above, there is another, physical reason for that, the **avalanche size exponent** τ. In traditional critical phenomena, the corresponding exponent of the order parameter distribution is fixed at unity in the presence of the Rushbrooke and the Josephson scaling law (Christensen *et al.* 2008). The deviation of τ from unity, which implies that the expected event size does not scale like

[5] The basic example $\mathcal{P}(s) = s^{-1}\mathcal{G}(s/s_c)$ with $\mathcal{G}(x) = 2x\exp(-x^2)/\sqrt{\pi}$ is normalised and has avalanche size exponent $\tau = 1$, as defined in Eq. (7.3). *Without* the pre-factor x in $\mathcal{G}(x)$ the graph looks surprisingly similar to a PDF as typically found in SOC models.

the characteristic event size, is another distinctive feature of SOC. To some extent, the exponent τ can be extracted from the avalanche size distribution (almost) by inspection. In a moment analysis, on the other hand, it is somewhat "hidden" in the details.

The most important observables usually extracted from an SOC model are thus the scaling exponents, such as τ, D (**avalanche dimension**), α (**avalanche duration exponent**) and z (**dynamical exponent**) discussed below. Here, the two exponents D and z are generally regarded as more universal than τ and α, as the former is often "enslaved" by an exact scaling law related to the average avalanche size, and the latter by a similar scaling law based on the "narrow joint distribution assumption", discussed in Sec. 7.2. Generally, all observables that are universal or suspected to be are of interest. This includes the scaling function (Sec. 7.2) which is most easily characterised by moment ratios, corresponding to universal amplitude ratios, traditionally studied in equilibrium critical phenomena (Privman *et al.* 1991, Salas and Sokal 2000).

7.1.2 Models

There is wide consensus on a number of general features of SOC models which seem to play a rôle in determining the universality class each belongs to. The very first SOC model, the BTW model, was essentially **deterministic**, *i.e.* there was no randomness in the bulk relaxation. A given configuration plus the site being charged next determines the resulting configuration uniquely. Even in these models, however, there can be a degree of **stochasticity**, namely when the site to be charged by the external drive is chosen at random. Finally, even when this is not the case, *i.e.* external drive and internal relaxation are deterministic, initial conditions are often chosen at random and averaged over.

Deterministic SOC models have the great appeal that they are "autonomous" (in a non-technical sense) or "self-sufficient" in that they do not require an additional source of (uncorrelated) noise. It is difficult to justify the existence of an external source which produces white, Gaussian noise, as that noise correlator, $\langle\eta(t)\eta(t')\rangle = 2\Gamma^2\delta(t-t')$, itself displays a form of scaling $\langle\eta(\alpha t)\eta(\alpha t')\rangle = \alpha^{-1}\langle\eta(t)\eta(t')\rangle$. The presence of an external (scaling) noise source seems to demote an SOC model to a conversion mechanism of scale invariance, which becomes most apparent when the respective model is cast in the language of stochastic equations of motion, *i.e.* **Langevin equations**.

Famous examples of deterministic SOC models, which do not require an external noise source for the relaxation process, are the BTW model with deterministic drive (Bak *et al.* 1987, but Creutz 2004), the OFC model (Olami *et al.* 1992) and, closely related, the train model (de Sousa Vieira 1992).

Of these only the latter has been studied extensively in the absence of all stochasticity.

Most SOC models, however, have a strong stochastic component, *i.e.* there is some randomness in the relaxation mechanism that gives rise to avalanches. In fact, models with some form of built-in randomness seem to give cleaner scaling behaviour, suggesting that deterministic models get "stuck" on some trajectory on phase space, where some conservation law prevents them from exploring the rest of phase space (Bagnoli *et al.* 2003, Casartelli *et al.* 2006). Notably, randomising the BTW model seems to push it into the Manna universality class (Karmakar *et al.* 2005). The latter model is probably the simplest SOC model displaying the most robust and universal scaling behaviour (Huynh *et al.* 2011). Due to the noise, trajectories of particles deposited by the external drive are those of random walkers.

The second dividing line distinguishes **Abelian** and **non-Abelian** models. The term was coined by Dhar (1990) introducing, strictly speaking, the Abelian Sandpile Model, by re-expressing the original BTW Model (Bak *et al.* 1987) in terms of units of slope rather than local particle numbers. This convenient choice of driving and boundary conditions renders the model unphysical as entire rows of particles are added and removed at once. At the same time, however, the model's final state after two consecutive charges at two different sites becomes independent from the order in which the charges and the subsequent relaxations are carried out. Practically all analytical insight into the BTW model is based on Dhar's (1990) Abelian version. Because it is easier to implement, it has also favoured in numerical simulations.

The term "Abelian" seems to suggest the existence of a (commutative) group, *i.e.* a set of operators closed under consecutive application, associative and containing inverse and an identity. For most SOC models referred to as Abelian, no such group is known, for example because operators do not exist explicitly, or the associative property makes little sense, similarly for the identity. Crucially, inverse operators rarely exist. To label a model Abelian therefore normally means that the final state does not depend on the order in which external charges are applied, *i.e.* the model updating operators (whether or not they exist), which drive it at various locations, commute. Because the final state is unique only in the case of deterministic models, stochastic models are Abelian provided that the statistics of the final state does not depend on the order in which external charges are applied (Dhar 1999b). The operators, which generally depend on the site the driving is applied to, of deterministic models apply to a model's state and take it from one quiescent state to the next. The operators in a stochastic model act on the distribution of states, *i.e.* they are the Markov operators. A deterministic model can be cast in the same language, however, the Markov operators then correspond to simple permutation matrices.

While Abelianness originally refers to the evolution of a model on the macroscopic time scale, it is generally used to characterise its behaviour on the microscopic timescale, *i.e.* the step-by-step, toppling-to-toppling update.

It is therefore usually concluded that the properties of avalanches and their statistics is independent from the order of toppling of multiple active sites.

Strictly, however, the Abelian symmetry does not apply to the microscopic time scale, at least for two reasons. Firstly, the Abelian operators apply, a priori, only to the avalanche-to-avalanche evolution, *i.e.* the macroscopic time scale. What is more, they apply to the final state and its statistics, but not necessarily to the observables. Applying charges at two different sites of an Abelian SOC model, starting from the same configuration, results in the same final state (or its statistics) regardless of the order in which the charges were applied, but not necessarily in the same pair of avalanche sizes produced. On the basis of the proof of Abelianness, at least in deterministic models, this limitation is alleviated by the insight that the sum of the avalanche sizes is invariant under a change of the order in which the model is charged.

As for the second reason, many models come with a detailed prescription of the microscopic updating procedure and therefore the microscopic time scale. Strictly, the invariance under a change of order of updates on the microscopic time scale thus applies to different models. The situation corresponds to equating different dynamics in the Ising model: For some observables, Glauber dynamics is different from Heat Bath dynamics, yet both certainly produce the same critical behaviour. In fact, choosing different dynamics (and thereby possibly introducing new conserved symmetries) can lead to different dynamical critical behaviour.

Revisting the proof of Abelianness, however, generally reveals that the caveats above are overcautious. The very proof of Abelianness on the macroscopic time scale uses and develops a notion of Abelianness on the microscopic time scale. This connection can be made more formally, once it has been established that any configuration, quiescent or not, can be expressed by applying a suitable number of external charges on each site of an empty lattice.

Abelianness generally plays a major rôle in the analytical treatment of SOC models, because it allows significant algebraic simplifications, not least when the dynamics of a model is written in terms of Markov matrices. It applies, generally, equally to recurrent and transient states, where no inverse exists. It remains highly desirable to demonstrate Abelianness on the basis of the algebra, once that is established as a suitable representation of a model's dynamics.

In the following section a few paradigmatic models of SOC are introduced: The BTW Model, the Manna Model, the OFC Model and the Forest Fire Model.

7.1.2.1 The BTW Model

The BTW Model was introduced together with the very concept of SOC (Bak *et al.* 1987), initially to explain the "ubiquity" of $1/f$ noise. Of course, since

then, SOC has been studied very much in its own right. Like virtually all SOC models, the BTW Model consists of a set of rules that prescribe how a local degree of freedom z_i on a d-dimensional lattice with sites i is to be updated. There are two different stages, namely the relaxation and the driving, the latter considered to be slow compared to the relaxation, *i.e.* the relaxation generally is instantaneous and never occurs simultaneously with the driving (separation of time scales). In the Abelian version of the BTW Model (Dhar 1990), the driving consists of adding a single slope unit (Kadanoff *et al.* 1989) to a site, that is normally picked uniformly and at random. The lattice is often initialised with $z_i = 0$ for all i.

If the driving leads to any of the z_i exceeding the the critical slope z^c (also referred to as the critical height or threshold, depending on the view) at a site i a **toppling** occurs whereby z_i is reduced by the coordination number q of the site and z_j of every nearest neighbour j increases by one (sometimes referred to as **charging**). In principle both q and z^c can vary from site to site and such generalisations are trivial to implement. It is common to choose $z^c = q - 1$.

The rules of the BTW Model can be summarised as follows:

Initialisation: All sites i are empty, $z_i = 0$.
Driving: One unit is added at a randomly chosen (or sometimes fixed) site i, *i.e.* $z_i \to z_i + 1$.
Toppling: A site with $z_i > z^c = q - 1$ (called *active*) distributes one unit to the q nearest neighbouring sites j, so that $z_i \to z_i - q$ and $z_j \to z_j + 1$.
Dissipation: Units are lost at boundaries, where toppling site i loses q units, $z_i \to z_i - q$, yet less than q nearest neighbours exist, which receive a unit.
Time progression: Time progresses by one unit per parallel update, when all active sites are updated at once.

A toppling can trigger an avalanche, as charged neighbours might exceed the threshold in turn, possibly by more than one unit. Strictly, the BTW Model is updated in parallel, all sites topple at once whose local degree of freedom exceeds the threshold at the beginning of a time step. Microscopic time then advances by one unit. This way, z_i might increase far beyond z^c before toppling itself. As long as $z_i > z^c$ for any site i, the sites in the model carry on toppling. The totality of the toppling events is an avalanche. In the Abelian BTW model as refined by Dhar (1990), the final state of the model does not depend on the order in which external charges are applied. In the process of the proof of this property, it turns out that the order of processing any charges during the course of an avalanche neither affects the final state nor the size of the avalanche triggered. Using a parallel updating scheme or not therefore does not change the avalanche sizes recorded. As the order of updates defines the microscopic time scale, a change in the updating procedure, however, affects all observables dependent on that time, such as avalanche duration or correlations on the fast time scale.

To keep the prescription above consistent with the notion of boundary sites, where toppling particles are to be lost to the outside, boundary sites have to be thought of as having the same number of nearest neighbours as any other, equivalent site in the bulk, except that some of their neighbours are not capable of toppling themselves. For numerical purposes it is often advisable to embed a lattice in some "padding" (a neighbourhood's "halo", see Sec. 7.3.2.2, p. 292), *i.e.* sites that cannot topple but are otherwise identical to all other sites.

The sum of the slope units residing on a given site i and those residing on its nearest neighbours remains unchanged by the toppling of site i, *i.e.* the bulk dynamics in the BTW are conservative. Dissipation occurs exclusively at the boundary and every slope unit added to the system in the bulk must be transported to the boundary in order to leave the system.

The original version of the BTW model is defined in terms of local heights, so that the height differences give rise to the slope z_i, which has to reach q in order to trigger an a toppling. While this is a perfectly isomorphic view of the BTW, *driving* it in terms of height units has a number of unwanted implications. In particular, it loses its Abelianness. For that reason, the original version of the BTW is rarely studied numerically nowadays.

The BTW Model is **deterministic** apart from the driving, which can be made deterministic as well, simply by fixing the site that receives the external charge that triggers the next avalanche. Even when slope units do not move independently at toppling, a randomly chosen slope unit being transported through a BTW system describes the trajectory of a random walker trajectories (Dhar 1990), essentially because every possible path is being realised (just not independently, but all with the correct weight). As a result, the average avalanche size $\langle s\rangle$ can be calculated exactly; The number of moves a slope unit makes on average from the time of being added by the external drive to the time it leaves the system through an open boundary is equal to the expected number of charges it causes. The expected number of charges (caused by the movement of all slope units taking part in an avalanche) per slope unit added is thus exactly equal to the expected number of moves a slope unit makes until it leaves the system, *i.e.* its escape time. If the avalanche size is measured by the number of topplings, which is more common, the expected number of moves has to be divided by the number of moves per toppling, q in the present case. Higher moments of the avalanche size, or, say, the avalanche size conditional to non-zero size (*i.e.* at least one site toppling in every avalanche), cannot be determined using the random walker approach, as they are crucially dependent on the interaction of toppling sites.

Due to the random walker property of the slope units added, the scaling of the average avalanche size thus merely depends on the particularities of the driving. If the driving is random and uniform, then $\langle s\rangle \propto L^2$ for any d-dimensional hypercubic lattice and (Ruelle and Sen 1992)

$$\langle s\rangle = \frac{1}{12}(L+1)(L+2) \tag{7.1}$$

in one dimension with two open boundaries, where the avalanche size is the number of topplings per particle added. However, the dynamics of the BTW Model in one dimension is trivial, so that the model is usually studied only in $d = 2$ and beyond.

Because (or despite of) its deterministic nature, a large number of analytical results are known, in one dimension (Ruelle and Sen 1992) but more importantly in two dimensions (Majumdar and Dhar 1992), not least on the basis of (logarithmic) conformal field theory (e.g. Majumdar and Dhar 1992, Ivashkevich 1994, Mahieu and Ruelle 2001, Ruelle 2002, Jeng 2005). Unfortunately, to this day, the scaling of the avalanche size distribution in dimensions $d \geq 2$ remains somewhat unclear. Numerically, results are inconclusive, as different authors quote widely varying results for $d = 2$ (Vespignani and Zapperi 1995, Chessa *et al.* 1999a, Lin and Hu 2002, Bonachela 2008, e.g.), possibly due to logarithmic corrections (Manna 1990, Lübeck and Usadel 1997, Lübeck 2000)

A major insight into the *collective* dynamics of toppling sites was the decomposition of avalanches into **waves** (Ivashkevich *et al.* 1994), which was later used by Priezzhev *et al.* (1996) to conjecture $\tau = 6/5$ for the avalanche size exponent in two dimensions. No site in an avalanche can topple more often than the site at which the avalanche was triggered. Not allowing that first site to topple therefore stops the avalanche from progressing any further and each toppling of the first site thus defines a wave of toppling.

While the BTW Model has been crucial for the formation of the field of SOC as a whole, its poor convergence beyond one dimension has made it fall in popularity. One may argue that the determinism of the dynamics is to blame, as found in other models (Middleton and Tang 1995). Indeed, adding some stochasticity makes the BTW Model display the universal behaviour of the Manna Model discussed in the next section (Černák 2002, Černák 2006).

The exponents reported for the BTW Model vary greatly. In two dimensions, the value of τ found in various studies ranges from 1 (Bak *et al.* 1987) to 1.367 (Lin and Hu 2002) and that for D from 2.50(5) (De Menech *et al.* 1998) to 2.73(2) (Chessa *et al.* 1999a). Similarly α is reported from 1.16(3) (Bonachela 2008) to 1.480(11) (Lübeck and Usadel 1997) and z from 1.02(5) (De Menech and Stella 2000) to 1.52(2) (Chessa *et al.* 1999a). Using comparatively large system sizes, Dorn *et al.* (2001) found exponents that seem to vary systematically with the system size with little or no chance to identify an asymptotic value.

The first exactly solved SOC model was the Dhar-Ramaswamy Model (Dhar and Ramaswamy 1989) which is the **directed** variant of the BTW Model. The directedness means that during an individual avalanche, sites are never re-visted, which effectively suppresses spatial correlations. Random drive of the model results in a product state, where sites taking part in an avalanche form a "compact" patch (*i.e.* they have no holes), which is delimited by boundaries describing a random walk. The exponents in $d = d_\perp + 1$ dimensions are given analytically by $D = 1 + d_\perp/2$, $D(2 - \tau) = 1$, $z = 1$

and $D(\tau-1)=z(\alpha-1)$, which implies $\alpha=D$ and $\tau=2-1/D$ (Dhar and Ramaswamy 1989, Christensen 1992, Christensen and Olami 1993, Tadić and Dhar 1997, Kloster *et al.* 2001). For example, in $d=1+1$ dimensions (directed square lattice), exponents are $D=3/2$, $\tau=4/3$, $z=1$ and $\alpha=3/2$. Mean-field exponents apply at $d=2+1$ and above.

7.1.2.2 The Manna Model

The Manna (1991) Model was originally intended as a simplified version of the BTW Model but has since then acquired the status of the paradigmatic representative of the largest (and maybe the only) universality class in SOC, generally refered to as the Manna, Oslo (Christensen *et al.* 1996) or C-DP (conserved directed percolation, Rossi *et al.* 2000) universality class.

The Manna Model displays robust, clean critical behaviour in any dimension $d\geq 1$, characterised by non-trivial exponents below $d=4$ (Lübeck and Heger 2003b). Originally, it is defined as follows: The external drive adds particles at random chosen sites i, *i.e.* the local degree of freedom increases by one, $z_i\to z_i+1$. If a site exceeds the threshold of $z^c=1$ it topples, so that *all* its particles are redistributed to the nearest neighbours, which are chosen independently at random. After the toppling of site i, the local degree of freedom is therefore set to $z_i=0$, while the total increase of the z_j at the nearest neighbours j of i maintains conservation. Again, as in the BTW model, non-conservation at boundary sites can be thought of as been implemented by sites that never topple themselves.

Charging neighbours might push their local degree of freedom beyond the threshold and they might therefore topple in turn. When a site topples, all particles present there at the time of toppling are transferred to its neighbour (maybe to a single one) and it is therefore crucial to maintain the order of (parallel) updates. The model is thus non-Abelian. In fact, the notion of Abelianness was initially restricted to deterministic models (Milshtein *et al.* 1998). However, Dhar (1999a) introduced a version of the Manna Model which is Abelian in the sense that the statistics of the final state remains unchanged if two consecutive external charges (by the driving) are carried out in reverse order. In that version of the Manna Model, a toppling site redistributes only 2 of its particles, *i.e.* the number of particles redistributed at a toppling does not depend on z_i itself. The difference between the BTW Model and the Manna Model lies thus merely in the fact that only two particles are re-distributed when a site topples in the Manna Model (irrespective of the coordination number of the site) and that the receiving sites are are picked at random.

In summary, the rules of the Abelian Manna Model are:

Initialisation: All sites i are empty, $z_i=0$.
Driving: One unit is added at a randomly chosen (or sometimes fixed) site i, *i.e.* $z_i\to z_i+1$.

Toppling: A site with $z_i > z^c = 1$ (called *active*) distributes one unit to 2 randomly and independently chosen nearest neighbouring sites j, so that $z_i \to z_i - 2$ and $z_j \to z_j + 1$.

Dissipation: Units are lost at boundaries, where the randomly chosen nearest neighbour might be outside the system.

Time progression: Originally, time progresses by one unit per parallel update, when all active sites are updated at once.

That the scaling in one dimension is not as clean as in higher dimension may be caused by logarithmic corrections (Dickman and Campelo 2003). Nevertheless, it has been possible to extract consistent estimates for exponents in dimensions $d = 1$ to $d = 5$ (Lübeck and Heger 2003b, Huynh *et al.* 2011, Huynh and Pruessner 2012). Because some of its exponents are so similar to that of the directed percolation universality class (Janssen 1981, Grassberger 1982, Hinrichsen 2000) there remains some doubt whether the Manna Model really represents a universality class in its own right (Muñoz *et al.* 1999, Dickman *et al.* 2002). The problem is more pressing in the **fixed energy** version (Dickman *et al.* 1998, Vespignani *et al.* 1998) of the Manna Model (Basu *et al.* 2012), where dissipation at boundaries is switched off by closing them periodically, thereby studying the model at a fixed amount of particles. The term "fixed energy sandpile" was coined to stress the conserved nature of the relavent degree of freedom (which may be called "energy") and to suggest a similar distinction as in the change of ensemble from canonical to microcanonical. Bonachela and Muñoz (2007) suggested to study the model with different boundary conditions which have an impact on the Manna Model that is distinctly different from that on models in the directed percolation universality class.

Because of its fixed energy version, the Manna Model is frequently studied for its links to absorbing state (AS) phase transitions (Dickman *et al.* 1998, Vespignani *et al.* 1998, Hinrichsen 2000, Henkel *et al.* 2008). In fact, it has been suggested that SOC is due to the self-organisation to the critical point of such an AS phase transition (Tang and Bak 1988, Dickman *et al.* 1998, Vespignani *et al.* 1998), whereby strong activity leads to a reduction of particles by dissipation, making the system in-active, while quiescence leads to activity due to the external drive. One may argue that such a linear mechanism cannot produce the desired universal critical behaviour without finely tuning the relevant parameters (Pruessner and Peters 2006, 2008, Alava *et al.* 2008).

A number of theoretical results are available for the Manna Model (Vespignani *et al.* 1998, 2000, Rossi *et al.* 2000, van Wijland 2002, Ramasco *et al.* 2004), yet an ϵ-expansion (Le Doussal *et al.* 2002) for the Manna universality class is available only via the mapping (Paczuski and Boettcher 1996, Pruessner 2003) of the Oslo Model (Christensen *et al.* 1996), which is the same universality class (Nakanishi and Sneppen 1997) as the Manna Model, to the quenched Edwards-Wilkinson equation (Bruinsma and Aeppli 1984, Koplik and Levine 1985, Nattermann *et al.* 1992, Leschhorn *et al.* 1997). Quenched

noise and disorder are, however, notoriously difficult to handle analytically. It is thus highly desirable to develop a better theoretical understanding of the Manna Model in its own right, including its mechanism of self-organisation, and to derive an ϵ-expansion for its exponents.

Although the Manna Model is more frequently studied in one dimension, for comparison with the BTW Model above, the exponents listed in the following were determined numerically in two dimensions for the Abelian and the non-Abelian (original) variant of the Manna Model. For τ they range from 1.25(2) (Biham *et al.* 2001) to 1.28(2) (Manna 1991, Lübeck and Heger 2003a), for D from 2.54 (Ben-Hur and Biham 1996) to 2.764(10) (Lübeck 2000), for α from 1.47(10) (Manna 1991) to 1.50(3) (Chessa *et al.* 1999b, Lübeck and Heger 2003a) and for z from 1.49 (Ben-Hur and Biham 1996) to 1.57(4) (Alava and Muñoz 2002, Dickman *et al.* 2002), generally much more consistent than in the BTW Model.

As in the BTW Model, various directed variants of the Manna Model which are exactly solvable for similar reasons as in the deterministic case have been extensively studied (Pastor-Satorras and Vespignani 2000b,a, Hughes and Paczuski 2002, Pan *et al.* 2005, Jo and Ha 2008). They have been characterised in detail by Paczuski and Bassler (2000b) and related to the deterministic directed models by Bunzarova (2010). Exponents generally follow $D = 3/2 + d_\perp/4$, which can be interpreted as the diffusive exploration of a random environment. Again, correlations are suppressed as sites are never re-visited in the same avalanche. As in the deterministic case, $z = 1$ and $D(2 - \tau) = 1$ and $D(\tau - 1) = z(\alpha - 1)$ result in $D = \alpha$. In $d = 1 + 1$ exponents are $\tau = 10/7$, $D = 7/4$, $\alpha = 7/4$ and $z = 1$.

7.1.2.3 The Forest Fire Model

The Forest Fire Model has an interesting, slightly convoluted history. Two distinct versions exist, which share the crucial feature that the bulk dynamics is not conservative. In the original version introduced by Bak *et al.* (1990) sites i, most frequently organised in a (two-dimensional) square lattice with periodic boundary conditions, can be in one of three states $\sigma_i \in \{T, F, A\}$, corresponding to occupation by a **T**ree, by **F**ire or by **A**sh. As time t advances in discrete steps, the state changes cyclically under certain conditions: A **T**ree turns into **F**ire at time $t + 1$ if a nearest neighbouring site was on **F**ire at time t. In turn, a **F**ire at time t becomes **A**sh in time $t+1$, and a site covered in **A**sh at time t might become occupied by a **T**ree at time $t + 1$ due to a repeated Bernoulli trial with (small) probability p. Starting from a lattice covered in trees, a single site is set on fire and the system evolves under the rules described. The key observable is the number of sites on fire as a function of time.

Initialisation: All (many) sites i contain a tree (otherwise ash), $\sigma_i = T$, and (at least) one site is on fire, $\sigma_i = F$.

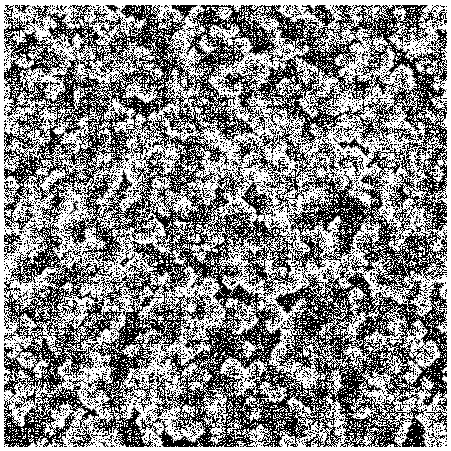


Fig. 7.1: Realisation of the original Forest Fire Model by Bak *et al.* (1990). **A**sh is marked by a white site, **T**rees are black and **F**ires grey.

Driving: With (small) probability p, a site i containing ash at the beginning of time step t contains a tree, $\sigma_i = A \to T$ at time $t + 1$.

Toppling: A site i that contains a tree at begining of time step t and has at least one nearest neighbour on fire, turns into fire as well, $\sigma_i = T \to F$. Simultaneously, a site on fire at t turns into ash, $\sigma_i = F \to A$.

Dissipation: trees grow slowly in Bernoulli trials and are removed in the "toppling". Their number is not conserved under any of the updating.

Time progression: Time progresses by one unit per parallel update.

The original Forest Fire Model (FFM) just described possesses an **absorbing state** from which it cannot recover within the rules given. If the fire stops spreading because the last site on fire is surrounded by ash, the only transition that can and will take place eventually occupies every site by a tree. Bak *et al.* (1990) originally suggested that occasional re-lightning might be necessary — in fact, if p is large enough, on sufficiently large lattices, there will always be tree to burn available. This, however, points to a fundamental shortcoming, as quantified by Grassberger and Kantz (1991), namely that the lengthscale of the relevant features of the FFM are determined by p. Typically, at small p, some large spiral(s) of fire keeps sweeping across the lattice. If p is chosen too small, the spatial extent of the spiral becomes too large compared to the size of the lattice and the fire eventually goes out. However, if a control parameter determines the characteristic length scale of the phenomenon, it cannot be *bona fide* SOC (e.g. Bonachela and Muñoz 2009). Figure 7.1 shows an example of the structures, most noticeable the fire fronts, developing.

The name "Forest Fire Model" should be taken as a witty aide-memoire. Bak *et al.* (1990) designed the model to understand scale free dissipation with uniform driving as observed in turbulent flow. The model should therefore be considered much more as a model of turbulence that happened to look like fires spreading in a forest. In the present model, perpetual fires spread across trees as they re-grow, which is a rather unrealistic picture; most fires in real forests are shaped by fire brigades, geographical and geological features and other environmental characteristics, as well as policies. Nevertheless, the original FFM as well as the version by Drossel and Schwabl (1992a), attracted significant attention as an actual model of forest fires, as well as other natural and sociological phenomena (Turcotte 1999).

There are two distinguishing features that set the FFM apart from many other SOC models. Firstly, the separation of time scales is incomplete, because driving the system by supplying new trees is a process running in parallel to the burning as fire spreads. Although the time scale of tree growth, parameterised by p, can in principle be made arbitrarily slow, the fire has to be constantly fed by new trees and cannot be allowed to go out, because there is no explicit re-lighting. In other words, the tree growth rates that still sustain fire are bounded from below. As a result, there are no distinct avalanches, as found in the BTW and the Manna Models.

More importantly, however, the FFM is different from other models because it is non-conservative at a fundamental level. No quantity is being transported to the boundaries and the local degree of freedom changes without any conservation.[6] At the time of the introduction of the FFM, it challenged Hwa and Kardar's (1989a) suggested mechanism of SOC that relied on a conservation law to explain the absence of a field-theoretic mass in the propagator.

Other dissipative models, like the SOC version of the "Game of Life" (Bak *et al.* 1989a), the OFC model discussed in the next section (Olami *et al.* 1992) and the Bak-Sneppen Model (Bak and Sneppen 1993) chipped away from the conservation argument put forward by Hwa and Kardar (1989a, 1992), Grinstein *et al.* (1990) and Socolar *et al.* (1993). The latter seem to have been caught by surprise by the advent of a variant of the FFM by Drossel and Schwabl (1992a) discussed in the following.

The Drossel-Schwabl Forest Fire Model (DS-FFM), as it is now normally referred to, was originally introduced by Henley (1989). It changes the original Forest Fire Model in two very important points: Firstly, the separation of time scales between burning and growing is completed, so that patches of (nearest neighbouring) trees are burned down instantly compared to all other

[6] It is difficult to make the statement about non-conservation strict. After all, the state of each site is meant to change and allowing for that, it is always possible to trace the appearance and the disappearance of something back to some influxes and outfluxes. Here is an attempt in the present case: While the increase in the number of trees can be thought of as being due to a corresponding influx, they can disappear with an enormous rate by spreading fire without explicit outflux *on that timescale*.

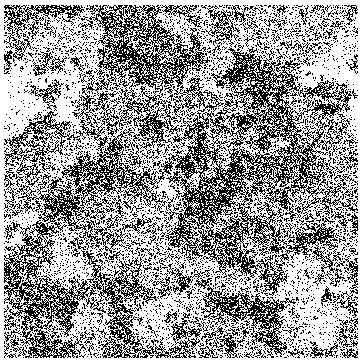


Fig. 7.2: Realisation of the Drossel-Schwabl Forest Fire Model (Drossel and Schwabl 1992a). **A**sh is marked by a white site, **T**rees are black.

processes. Because fires therefore burn down completely before anything else can happen, fires are set, secondly, explicitly by random, independent uniform lightning. The key-observables of the DS-FFM are the geometrical features of the clusters burned down, such as the number of occupied sites (the mass) and the radius of gyration.

While trees grow with rate p on every empty site (*i.e.* one containing ash), lightning strikes with much lower rate f on every site. If it contains a tree, the fire eradicates the entire cluster of trees connected to it by nearest neighbour interactions. In summary:

Initialisation: All sites i contain ash, $\sigma_i = A$.
Driving: With (small) probability p, a site i containing ash at the beginning of time step t contains a tree, $\sigma_i = A \to T$ at time $t + 1$.
Toppling: With probability $f \ll p$, a site containing a tree at the beginning of time step t and the entire cluster of trees connected to it by nearest neighbour interactions is changed to ash, $\sigma_i = T \to A$.
Dissipation: trees grow slowly in Bernoulli trials and are removed in the "toppling". They are not conserved in any of the updates.
Time progression: Time progresses by one unit per parallel update, toppling is instantaneous relative to growing trees.

As a result entire patches of forest disappear at a time, which are reforested with the same Poissonian density p. This process results in a patchy structure with individual islands having roughly homogeneous tree-density, Figure 7.2.

In a change of perspective, the processes parameterised by p and f are tree growth attempts and lightning attempts which fail if the site is already occupied by a tree or does not contain one, respectively. The original definition by Drossel and Schwabl (1992a) still used discrete time, so that both p and f were probabilities, rather than Poissonian rates, which can be recovered

by rescaling p and f simultaneously. However, it is common (e.g. Clar *et al.* 1996) to rescale time so that $p = 1$ (enforced growth on randomly picked empty sites) and to attempt p/f times to grow a tree before attempting to set one alight. In order to see scale-free cluster size distributions, a **second separation of timescales** is needed, whereby the ratio p/f diverges.

Many of the properties of the DS-FFM are percolation-like. If it were not for the correlations in the tree-density, which develop because of "synchronous, patchy re-forestation", i.e. if the tree-density was homogeneous, then the DS-FFM would be a form of percolation. In particular, the cluster size distribution (of the patches removed and the totality of all patches present) was given by that of (well-known) static percolation.

The DS-FFM does not suffer from the same short-coming as the original FFM of having a well-understood typical (spiral) structure, whose size is determined by the single control parameter p, yet it still has one control parameter which needs to be finely tuned in accordance with the system size. This parameter is p/f — if it is too large, then the lattice will be densely filled with trees before lightning strikes and removes almost all of them, leaving behind essentially a clean sheet with a few remaining (small) islands of trees. If p/f is too small, then no dense forest ever comes into existence and the cluster size distribution has a cutoff not determined by the system size, but by that parameter.

In extensive numerical studies (Grassberger 2002, Pruessner and Jensen 2002a, 2004), the system sizes were chosen big enough for each p/f that finite size effects were not visible, *i.e.* for each p/f convergence of the cluster size distribution $\mathcal{P}(s;L)$ in the system size L was achieved. However, these studies revealed that the DS-FFM does not display simple scaling in $s_c = s_c(p/f)$, Eq. (7.3) (Sec. 7.2.1). While $\mathcal{P}(s)/s^{-\tau}$ converges in the thermodynamic limit (as it should, trivially) for any τ, there is no choice of τ so that the remaining functional profile depends only on the ratio $s/s_c(p/f)$. Instead, $\mathcal{P}(s)/s^{-\tau}$ depends explicitly on both s and $s_c(p/f)$, or, for that matter, p/f. The only feature that may display some convergence (Pruessner and Jensen 2002a) is the bump in the probability density function (PDF) towards large s. For some choice of τ, there is a small region, say $[s_c(p/f)/2, s_c(p/f)]$, where $\mathcal{P}(s)/s^{-\tau}$ traces out a very similar graph, as if the lower cutoff s_0 itself was a divergent multiple of the upper cutoff.[7]

One may hope that finite size scaling can be recovered, taking the limit of large p/f and considering $\mathcal{P}(s)/s^{-\tau}$ as a function of L. However, it is clear that the PDF trivialises in this limit,

$$\lim_{p/f\to\infty} \mathcal{P}(s;p/f,L) = s^{-1}\delta\left(\frac{s}{L^d}\right) \tag{7.2}$$

[7] If $s_c(p/f)$ marks roughly the maximum of the bump, the PDF drops off beyond it so quickly, that next to nothing is known of $\mathcal{P}(s)$ beyond s_c. In principle, however, if there is approximate coincidence on $[s_c(p/f)/2, s_c(p/f)]$, there should also be approximate coincidence on $[s_c(p/f)/2, \infty)$.

as the lattice is completely covered in trees before they all get completely removed in a singly lightning.

Interestingly, the lack of scaling in finite $s_c(p/f)$ is not visible in the scaling of the moments $\langle s^n \rangle$ because they are sensitive to large event sizes (at any fixed $n > \tau - 1$), rather than the smaller ones around the lower cutoff, whose divergence violates simple scaling.

As in the BTW Model, exponents reported for the DS-FFM (if they are reported at all) display a fairly wide spread. In two dimensions, they are τ from 1 (Drossel and Schwabl 1992a) to 1.48 (Patzlaff and Trimper 1994) and D from 1 (Drossel and Schwabl 1992a) to 1.17(2) (Henley 1993, Honecker and Peschel 1997).

7.1.2.4 The OFC Model

To this day, the Olami-Feder-Christensen Model (OFC Model Olami *et al.* 1992) is one of the most popular and spectacular models of SOC. It is a simplified version of the Burridge-Knopoff Model (Burridge and Knopoff 1967) of earthquakes, it has a *tunable* degree of non-conservation (including a conservative limit) with a clear physical meaning, it has been extensively analysed, both in time and space, for the effect of different boundary conditions (Middleton and Tang 1995), and its one-dimensional variant (de Sousa Vieira 1992) has been linked to the Manna universality class (Paczuski and Boettcher 1996, Chianca *et al.* 2009). After the definition of the model, the discussion below focuses on the model's rôle in earthquake modelling and the attention it received for the spatio-temporal patterns it develops.

The OFC Model is at home on a two-dimensional square lattice. As in the models above, each site i has a local degree of freedom $z_i \in \mathbb{R}$ (called the local "pulling force"), which is, in contrast to the models above, however, real-valued. As in the BTW Model, there are two clearly distinct stages of external driving and internal relaxation. During the driving *all* sites in the system receive the same amount of force (sometimes referred to as "continuous" or better "uniform" drive) until one site exceeds the threshold $z^c = 1$, which triggers a relaxation during which no further external driving is applied. In a relaxation or toppling, a site re-distributes a fraction of *all* pulling force evenly among its nearest neighbours which may in turn exceed the threshold. The force z_i at the toppling site i is set to 0 and the amount arriving at each neighbour is αz_i, where α is the **level of conservation**. At coordination number q, a level conservation less than $1/q$ means that the bulk dynamics is dissipative. Boundary sites lose force αz_i (at corners multiples thereof) to the outside. Because the force re-distributed depends on the amount of pulling force present at the site at the time of the re-distribution, the order of updates matters greately, *i.e.* the OFC Model is not Abelian. If $\alpha < 1/q$ periodic boundary conditions can be applied without losing the possibility of a stationary state, yet normally the boundaries are open. The OFC Model is

normally initialised by assigning random, independent forces from a uniform distribution.

Sites to topple are identified at the beginning of a timestep and only those have been relaxed by the end of it (parallel updates). Unless more than one site exceeds the threshold (degenerate maximum) at the beginning of an avalanche, toppling sites therefore reside on either of the two next nearest neighbour sublattices of a square lattice.

Again, a separation of time scales is applied, where the relaxation becomes infinitely fast compared to an infinitesimally slow drive. In an actual implementation, however, the driving is applied instantaneously and the relaxation takes up most (computational time): The driving can be completed in a single step by keeping track of the site, say i^* with the largest pulling force acting on it. The amount of force added throughout the system is thus simply $z^c - z_{i^*}$, triggering the next avalanche.

Because sweeping the lattice in search of the maximum is computationally very costly,[8] the main computational task in the OFC Model is to keep track of the site exposed to the maximum pulling force. This is a classic computational problem (Cormen *et al.* 1996), which also is occurs in other models, such as the Bak-Sneppen Model (Bak and Sneppen 1993). The traditional solution is to organise data in a tree-like structure and devise methods that allow fast updating and determination of the maximum. However, in the OFC Model updating as site's force is much more frequent than determination of the maximum and thus a fast algorithm focuses on the optimisation of the former at the expense of the latter, *i.e.* a slightly slower procedure to determine the maximum.

Grassberger (1994) pointed out a number of improvements to a naïve, direct implementation of the OFC Model. Firstly, instead of driving the system uniformly, thereby having to sweep the lattice to increase the force on every site by $z^c - z_{i^*}$, the threshold z^c is to be lowered; the amount of force re-distributed at toppling is obviously to be adjusted according to the new offset. The second major improvement Grassberger (1994) suggested was the organisation of forces in "boxes" (sometimes referred as **Grassberger's box-technique**), which splits the range of forces present in the system in small enough intervals that the search for the maximum force succeeds very quickly, yet keeps the computational effort to a minimum when re-assigning a box after an update. Other improvements suggested was maintaining a stack (Sec. 7.3.1) of active sites, and the use of a scheme to determine neighbouring sites suitable to the programming language at hand.

The adjustment of z^c outlined above has some rather unexpected effects depending on the numerical precision (Sec. 7.3.3) used in the simulation (Pruessner 2012b). As pointed out by Drossel (2002), the OFC Model is ex-

[8] Not only is the very searching *across all sites* costly, most of the memory occupied by the lattice will not reside in a cache line (as for example most "local" data) and thus has to be fetched through a comparatively slow bus.

tremely sensitive to a change of precision; a lower precision seems to enhance or favour phase-locking, discussed in the following.

Most of the studies of SOC models focuses on large-scale statistical features, large both in time and space. The analysis of the OFC Model by Socolar *et al.* (1993) Middleton and Tang (1995) and Grassberger (1995) therefore ventured into uncharted territory as they studied the evolution towards stationarity in the OFC Model on a microscopic scale, analysing the patchy structure of the forces on the lattice.

Firstly, periodic boundary conditions inevitably lead to periodic behaviour in time. Below $\alpha \approx 0.18$ in a two-dimensional square lattice, (almost) every avalanche has size unity. In that extreme case, the period is strictly $1 - q\alpha$, because discounting the external drive, this is the amount of force lost from every site after every site has toppled exactly once, as the system goes through one full period.

The periodicity is broken once open boundaries are introduced. Sites at the edge of the lattice have fewer neighbours that charge them, so if every site in the system topples precisely once, the force acting on a boundary site is expected to be lower. While open boundaries indeed break temporal periodicity, they form, at the same time, seeds for (partially) synchronised patches, which seem to grow from the outside towards the inside, increasing in size. Middleton and Tang (1995) introduced the term **marginal (phase) locking** to describe this phenomenon.

The temporal periodicity might similarly be broken by introducing inhomogeneities or disorder, effective even at very low levels (Grassberger 1994, Ceva 1995, 1998, Torvund and Frøyland 1995, Middleton and Tang 1995, Mousseau 1996). That a spatial inhomogeneity helps forming synchronised patches in space can also be attributed to marginal phase locking.

Because the OFC Model is so sensitive to even the smallest amount of disorder and inhomogeneity, its statistics is often taken from very big samples with extremely long transients. Many authors also average over the initial state. Drossel (2002) suggested that despite these precautions, some of the statistical behaviour allegedly displayed by the OFC Model might rather be caused by numerical "noise", also a form of inhomogeneity or disorder entering into a simulation. In practise, it is difficult to discriminate genuine OFC behaviour from numerical shortcomings and one may wonder whether some of these shortcomings are not also present in the natural phenomenon the OFC Model is based on.

That SOC may be applicable in seismology had been suggested by Bak *et al.* (1989b, also Bak and Tang 1989, Sornette and Sornette 1989, Ito and Matsuzaki 1990) at a very early stage. The breakthrough came with the OFC Model, which is based on the Burridge-Knopoff Model of earthquakes (or rather fracturing rocks). The latter is more difficult to handle numerically, with a "proper" equation of motion taking care of the loading due to spring-like interaction much more carefully. The OFC Model, on the other hand,

is much easier to update, almost like a cellular automaton.[9] The context of SOC provided an explanatory framework of the scale-free occurrence of earthquakes as described by the **Gutenberg-Richter** law (Gutenberg and Richter 1954, Olami *et al.* 1992). Even though exponents both in the real-world as well as in the OFC Model seem to lack universality, certain scaling concepts, motivated by studies in SOC, have been applied successfully to earthquake catalogues (Bak *et al.* 2002).

It is fair to say that the OFC Model, to this day, is widely disputed as a *bona fide* model of earthquakes. Its introduction has divided the seismology community, possibly because of the apparent disregard of their achievements by the proponents of SOC (Bak and Tang 1989). One of the central claims made initially is that earthquakes are unpredictable if they are "caused" by SOC, which questions the very merit of seismology. However, given that SOC is a framework for the understanding of natural phenomena on a long time and length scale, providing a mechanism for the existence of long temporal correlations, SOC indicates precisely the opposite of unpredictability. This point is discussed controversially in the literature to this day (Corral 2003, 2004c,b, Davidsen and Paczuski 2005, Lindman *et al.* 2005, Corral and Christensen 2006, Lindman *et al.* 2006, Werner and Sornette 2007, Davidsen and Paczuski 2007, Sornette and Werner 2009). Older reviews (Turcotte 1993, Carlson *et al.* 1994) help to understand the historical development of the dispute. Hergarten (2002) and more recently Sornette and Werner (2009) have put some of the issues in perspective.

There is not a single set of exponents for the OFC Model, as they are generally expected to vary with the level of conservation (Christensen and Olami 1992). Because authors generally do not agree on the precise value of α to focus on, results are not easily comparable across studies. Even in the conservative limit, $\alpha = 1/q$, little data is available, suggesting $\tau = 1.22(5) - 1.253$ and $D = 3.3(1) - 3.01$ (Christensen and Olami 1992, Christensen and Moloney 2005).

7.2 Scaling and numerics

As a rule of thumb, SOC models are **SDIDT** systems Jensen (1998): **S**lowly **D**riven **I**nteraction **D**ominated **T**hreshold systems. The driving implements a separation of time scales and thresholds lead to highly non-linear interaction, which results in avalanche-like dynamics, the statistics of which displays scaling, a continuous symmetry. Ideally, the scaling behaviour of an SOC model can be related to some underlying continuous phase transition, which is triggered by the system self-organising to the critical point.

[9] Strictly, the OFC Model generally is not a cellular automaton, because the local states z_i are continuous.

The critical behaviour can be characterised by (supposedly) universal critical exponents, the determination of which is the central theme of the present section. At the time of the conception of SOC, critical exponents were extracted directly from probability density function, (PDFs), often by fitting the data to a straight line in double-logarithmic plot. Frequently, such scaling is referred to as "power law behaviour". Very much to the detriment of the entire field, some authors restrict their research to the question whether an observable displays the desired behaviour, without attempting to determine its origin and without considering the consequences of such behaviour. Power law behaviour therefore has become, in some areas, a mere curiosity.

7.2.1 Simple scaling

While studying power laws in PDFs can be instructive, there are far superior methods to quantify scaling behaviour. In recent years, most authors have focused on an analysis of the moments of the PDFs, as traditionally done in the study of equilibrium statistical mechanics. Not only is this approach more efficient, it also is more accurate and mathematically better controlled. Moreover, it is concerned directly with an observable (or rather, arithmetic means of its powers), rather than its accumulated histogram.

Nevertheless the starting point of a scaling analysis in SOC, is the **simple (finite size) scaling assumption**,

$$\mathcal{P}(s) = as^{-\tau}\mathcal{G}(s/s_c) \text{ for } s \gg s_0 \ , \tag{7.3}$$

where $\mathcal{P}(s)$ is the (normalised) probability density function of an observable, s in this case, a is a (non-universal) **metric factor** present to restore dimensional consistency and accounting for the (microscopic) details of the model, τ is a **universal scaling** (or critical) **exponent**, $\mathcal{G}$ is a **universal scaling function**, s_c is the **upper cutoff** and s_0 the **lower cutoff**. If s is the avalanche size, then τ is known as the **avalanche size exponent**, when s is the duration, then τ is traditionally replaced by α and called the **avalanche duration exponent**.

The two cutoffs signal the onset of new physics: Below s_0 some microscopic physics prevails, often a lattice spacing or some other minimal length below which discretisation effects take over. Above s_c some large finite length scale becomes visible, which in SOC is normally controlled by the size of the lattice, so that Eq. (7.3) is referred to as *finite size* scaling. In traditional critical phenomena, s_c is controlled by the correlation length, beyond which distant parts of the system can be thought of as being independent, suggesting the validity of the central limit theorem.

Strictly, SOC models should always tune themselves to a critical point, so that the algebraic, critical behaviour is cut off only by the system size.

All scaling in SOC therefore is finite size scaling. There are a handful of established SOC models, which violate that strict rule, however, such as the Drossel-Schwabl Forest Fire Model Drossel and Schwabl (1992a), where an additional parameter has to be tuned simultaneously with the system size.

The physical origin of the scales contained in the metric factor a and the lower cutoff s_0 often is the same, yet even with these length scales present, $\mathcal{P}(s)$ has an arbitrarily wide region where it displays a power-law dependence on s and whose width is controlled by s_c; if $s_0 \ll s \ll s_c$, then $\mathcal{P}(s) = as^{-\tau+\alpha}s_c^{-\alpha}\tilde{\mathcal{G}}_0$, provided

$$\lim_{x\to 0} x^{-\alpha}\mathcal{G}(x) = \tilde{\mathcal{G}}_0 \ . \tag{7.4}$$

Typically, however, $\alpha = 0$ so that the intermediate region of $\mathcal{P}(s)$ displays a power law dependence with exponent τ, which can in principle be extracted as the negative slope of $\mathcal{P}(s)$ in a double logarithmic plot. However, because it is *a priori* unclear whether the scaling function $\mathcal{G}(s/s_c)$ can be approximated sufficiently well by a constant $\mathcal{G}_0$, "measuring" the exponent τ by fitting the intermediate region of a double logarithmic plot to a straight line (sometimes referred to as the **apparent exponent**) is very unreliable. If the scaling function displays a power law dependence on the argument, $\alpha \neq 0$, the effective exponent in the intermediate region is $\tau - \alpha$. One can show that α is non-negative, $\alpha \geq 0$, and $\tau = 1$ if $\alpha > 0$ (Christensen *et al.* 2008).

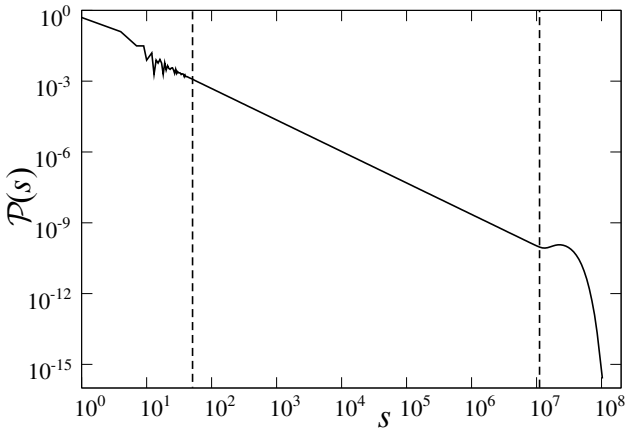


Fig. 7.3: Example of a double logarithmic plot of the PDF of the avalanche size in an SOC model (Data from Pruessner 2012b).

Figure 7.3 shows a typical double-logarithmic plot of the PDF in an SOC model. The power law region is marked by two dashed lines. The lower cutoff is at around $s_0 = 50$ and the features below that value are expected to be essentially reproduced by that model irrespective of its upper cutoff. The

spiky structure visible in that region is not noise and may, to some extent, be accessible analytically, similar to the lattice animals known in percolation (Stauffer and Aharony 1994). The power law region between the two dashed lines can be widened arbitrarily far by increasing the upper cutoff s_c. Running the same model with increasing s_c will reproduce this almost straight region beyond which the bump in the data indicates the onset of the upper cutoff.

The upper cutoff in SOC models supposedly depends only on the system size and does so in a power-law fashion itself,

$$s_c(L) = bL^D \tag{7.5}$$

where b is another metric factor and D is the **avalanche dimension**. The exponent describing the same behaviour for the upper cutoff of the avalanche duration is the **dynamical exponent** z. The four exponents τ, D, α and z are those most frequently quoted as the result of a numerical study of an SOC model.

The simple scaling ansatz Eq. (7.3) as well the scaling of the upper cutoff, Eq. (7.5), both describe *asymptotic* behaviour in large s_c and L respectively. When determining exponents in computer simulations of SOC models, corrections have to be taken into account in a systematic manner. While subleading terms are difficult to add to the simple scaling ansatz Eq. (7.3), this is routinely done in the case of the upper cutoff,

$$s_c(L) = bL^D(1 + c_1L^{-\omega_1} + c_2L^{-\omega_2} \ldots) \tag{7.6}$$

Corrections of this form are referred to as **corrections to scaling** (Wegner 1972) or confluent singularities. These corrections are discussed further in the context of moment analysis, Sec. 7.2.2.

Although some very successful methods of analysis exist (Clauset *et al.* 2009), Eq. (7.3) does not lend itself naturally to a systematic quantitative analysis for fixed s_c. Often, a **data collapse** is performed in order to demonstrate the consistency of the data with simple scaling. According to Eq. (7.3) the PDF $\mathcal{P}(s)$ for different cutoffs s_c produces the same graph by suitable rescaling, in particular by plotting $\mathcal{P}(s)\, s^\tau$ against $x = s/s_c$, which gives $\mathcal{G}(x)$. Deviations are expected for small values of s/s_c, namely for s around s_0, where Eq. (7.3) does not apply. Figure 7.4 shows an example of such a collapse using the same model as in Figure 7.3.

Provided $\lim_{x\to 0} \mathcal{G}(x) = \mathcal{G}_0 \neq 0$, the region where $\mathcal{P}(s)$ displays (almost) a power law translates into a horizontal, (nearly) constant section in the rescaled plot. The graph terminates in a characteristic **bump**, where the probability density of some larger event sizes exceeds that of some large, but smaller ones. This counter-intuitive feature is normally interpreted as being caused by system spanning events which were terminated prematurely by the boundaries of the system. Had the system been larger, the events would have developed further. In the PDF of a larger system thus make up the regular, straight power law region, where the smaller system's PDF displays a bump.

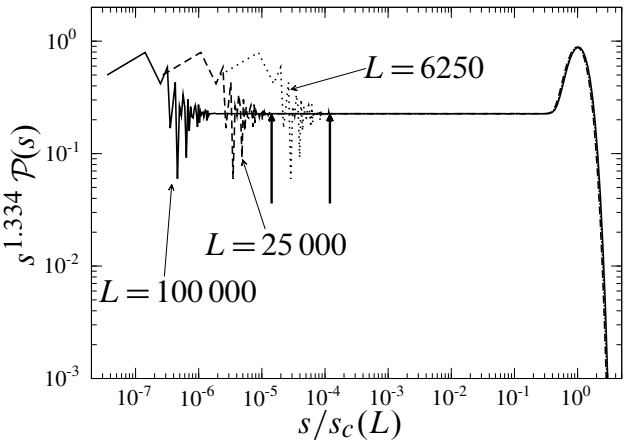


Fig. 7.4: Data collapse of three different data sets similar to the data shown in Figure 7.3. The upper cutoff s_c is solely controlled by the system size L (Data from Pruessner 2012b).

Even when the total probability contained in the bump is finite but very small, it is enough to account for all events contained beyond it in the power law region of an infinite system.

A data collapse is not unique, as plotting $\mathcal{P}(s)\, s^\tau f(s/s_c)$ produces $\mathcal{G}(x)f(x)$ for any function $f(x)$. In the literature, $f(x)$ is often chosen as $f(x) = x^{-\tau}$ so that $\mathcal{P}(s)\, s^\tau f(s/s_c) = \mathcal{P}(s)\, s_c^\tau$. Plotting that data has the fundamental disadvantage that $\mathcal{P}(s)\, s_c^\tau$ usually spans many orders of magnitude more across the ordinate compared to $\mathcal{P}(s)\, s^\tau$, so that details in the terminal bump are less well resolved.

7.2.1.1 Binning

A clean, clear dataset like the one shown in Figure 7.3 is the result of **binning**. For numerical studies of SOC models this is a necessary procedure in order to smoothen otherwise rather rugged histograms. The reason for that ruggedness is the strong dependence of the probability density on the event size, with very few large events occurring. Because of the power law relationship between event size and frequency, their total numbers decrease even on a logarithmic scale. As a result, statistical noise visibly takes over, often clearly before the onset of the terminal bump. Statistics for large event size is sparse and often little more than a muddle of seemingly unrelated data points is visible in the raw data for large events.

The noise can be reduced by averaging the data for increasingly large event sizes over increasingly large "bins", hence the name binning. This is

normally done in **post-processing** of the raw data produced in a numerical simulation, by summing over all events within a bin and dividing by its size. In principle, the bin sizes could be chosen to fit the data; if the bin ranges are $[b_i, b_{i+1})$, then a pure power law $\mathcal{P}(s) = as^{-\tau}$ would deposit

$$\int_{b_i}^{b_{i+1}} \mathrm{d}s\, as^{-\tau} = \frac{a}{\tau - 1} \left(b_i^{1-\tau} - b_{i+1}^{1-\tau} \right) \tag{7.7}$$

events in each bin i. This number can be made constant by choosing $b_i = (B_0 - B_1 i)^{1/(1-\tau)}$. Similarly, one might chose the bin boundaries b_i "on the fly", *i.e.* successively increase the bin size until roughly a given number of entries have been collected. While those two choices lead to uniformly low statistical errors (assuming constant correlations), they both suffer from significant shortcomings. Firstly, the exponent τ to be estimated from the data should not enter into the very preparation of the data that is meant to produce the estimate. This problem is mitigated by the fact that τ may be determined through a separate, independent procedure. Secondly and more importantly, both procedures will lead to an increasingly wide spacing of data points, which becomes unacceptable towards large event sizes, because the abscissa will no longer be defined well enough — if b_{i+1} and b_i are orders of magnitude apart, which s does Eq. (7.7) estimate. Last but not least, to make PDFs of different system sizes comparable, the same b_i should be used for all datasets.

The widely accepted method of choice is **exponential binning** (sometimes also referred to as **logarithmic binning**), where $b_i = B_0 \exp(\beta i)$. Such bins are equally spaced on the abscissa of a double logarithmic plot. Because the width of exponential bins is proportional[10] to their limits, Eq. (7.7), sparse data can cause a surprising artefact, whereby single events spuriously produce a probability density which decays inversely with the event size, $\mathcal{P}(s) \propto s^{-1}$, suggesting an exponent of $\tau = 1$. A typical problem with exponential bins occurs at the small end of the range when used for integer valued event sizes, because in that case the $b_{i+1} - b_i$ should not be less than 1. It is then difficult to control the number of bins and thus the resolution effectively, because decreasing β increases the number of minimally sized bins and has highly non-linear knock-on effects on all bin boundaries. The problem is obviously much less relevant for non-integer event sizes, such as the avalanche duration. However, it is rather confusing to use non-integer bin boundaries for integer valued event sizes, because bins may remain empty and the effective bin size cannot be derived from $b_{i+1} - b_i$. For example a bin spanning $b_{i+1} - b_i = 0.9$ may not contain a single integer, whereas $b_{i+1} - b_i = 1.1$ may contain two.

It is obviously advantageous to perform as much as possible of the data manipulation as post-processing of raw simulation data. Efficiency and memory limitations, however, normally require a certain level of binning at the

[10] For integer valued bin boundaries, strictly, this holds only approximately.

simulation stage. When event sizes and frequencies spread over 10 orders of magnitude a simple line of code[11]

```
histogram[size]++; /* one count for size in the histogram */
```

would require `histogram` to have a precision of more than 32 bits. Normally such counters are implemented as integers, which would need to be a `long long int` in the present case. The memory required for 10^{10} of these 64 bit numbers (about 75 GB) exceeds by far the memory typically available in computers in common use at the time of writing this text (2012). Writing every event size in a list, eventually to be stored in a file, is rarely an alternative, again because of the enormous memory requirements and because of the significant amount of computational time post-processing would take.

Consequently, some form of binning must take place at the time of the simulation. In principle, any sophisticated binning method as used during post-processing can be deployed within the simulation, yet the risk of coding errors ruining the final result and the computational effort renders this approach unfeasible. The established view that complicated floating point operations such as `log` or `pow` are too expensive to be used regularly in the course of a numerical simulation has experienced some revision over the last decade or so, as techniques like hyperthreading and out-of-order execution are commonly used even in the FPU. Nevertheless, integer manipulation, often doable within a single CPU cycle, remains computationally superior compared to floating point manipulation. Even some of the rather archaic rules remain valid, such as multiplications being computationally more efficient than divisions, as they can be performed within a short, fixed number of cycles. Further details can be found in the appendix at the end of the chapter.

7.2.2 Moment analysis

By far the most powerful technique to extract universal features of an SOC model is a moment analysis (De Menech *et al.* 1998). Traditionally, the numerical investigation of critical phenomena has focused on moments much more than on the underlying PDF, even when the former are often seen as the "result" of the latter. Mathematically, no such primacy exists and one can be derived from the other under rather general conditions (Feller 1966, Carleman's theorem in). In general one expects that a finite system produces only finite event sizes, *i.e.* that finite systems have a sharp cutoff of the "largest possible event size". While very physical, this rule finds its exception in residence times, when particles get "buried" in a "pile" over long periods. In the

[11] All explicit examples in this chapter are written in C, which is the most widely used programming language for numerical simulations, as long as they are not based on historic Fortran code.

Oslo Model, some of these waiting time distributions seem to be moderated by scaling functions that are themselves power laws and may possess upper cutoffs exponential in the system size (Dhar and Pradhan 2004, Pradhan and Dhar 2006, 2007, Dhar 2006).

Assuming, however, that all moments

$$\langle s^n \rangle = \int_0^\infty \mathrm{d}s \, s^n \mathcal{P}(s) \tag{7.8}$$

exist, *i.e.* are finite, then for $n + 1 > \tau$

$$\langle s^n \rangle \simeq a s_c^{1+n-\tau} g_n \tag{7.9}$$

where $\simeq$ is used to indicate equivalence *to leading order in large* s_c. Moments with $n < \tau - 1$ are not dominated by the scaling in s_c, *i.e.* they are convergent in large s_c. The (asymptotic) amplitudes g_n are defined as

$$g_n = \int_0^\infty x^{n-\tau} \mathcal{G}(x) \tag{7.10}$$

expected to be finite for all $n \geq 0$. There is an unfortunate confusion in the literature about the (spurious) consequences of $\langle s^0 \rangle = 1$ scaling like $s_c^{1-\tau} g_0$. If $\tau > 1$, then the leading order of $\langle s^0 \rangle$ is not given by Eq. (7.9).

The only scaling in SOC is finite size scaling, *i.e.* the upper cutoff is expected to diverge with the system size, Eq. (7.5), so that moments scale like

$$\langle s^n \rangle \simeq a b^{1+n-\tau} L^{D(1+n-\tau)} g_n \ . \tag{7.11}$$

Neither a nor b are universal and neither are the g_n unless one fixes some features of $\mathcal{G}(x)$ such as its normalisation and its maximum. To extract universal characteristics of $\mathcal{G}(x)$, moment ratios can be taken for example

$$\frac{\langle s^{n-1} \rangle \langle s^{n+1} \rangle}{\langle s^n \rangle^2} = \frac{g_{n-1} g_{n+1}}{g_n^2} + \text{corrections} \tag{7.12}$$

or

$$\frac{\langle s^n \rangle \langle s \rangle^{n-2}}{\langle s^2 \rangle^{n-1}} = \frac{g_1^{n-2}}{g_2^{n-1}} g_n + \text{corrections} \ , \tag{7.13}$$

which is particularly convenient because of its very simple form when fixing $g_1 = g_2 = 1$ by choosing the metric factors a and b appropriately.

The most important result of a moment analysis, however, are the universal exponents D and τ and corresponding pairs for avalanche duration (z and α respectively), as well as the area (normally D_a and τ_a) *etc.*. This is done in a three step process. Firstly, the SOC model is run with different systems sizes, typically spaced by a factor 2, or $2, 5, 10$. It can pay to use slightly "incommensurate" system sizes to identify systematic effects, for example

due to boundary effects being particularly pronounced in system sizes that are powers of 2. A typical simulation campaign would encompass 10 to 15 system sizes, of which maybe only 6 to 10, stretching over two to four orders of magnitude[12] will be used to produce estimates of exponents. The result of the simulation are estimates for the moments of the relevant observables together with their error (see below).

Secondly, the moments of the event sizes distribution, $\langle s^n \rangle$, are fitted against a power law in L (which is the parameter controlling s_c) with corrections,

$$\langle s^n \rangle = A_0 L^{\mu_n} + A_1 L^{\mu_n - \omega_1} + \dots \tag{7.14}$$

with positive exponents ω_i, known as confluent singularities; in particular $\mu_n - \omega_1$ is sometimes referred to as a sub-dominant exponent. The introduction of such **corrections to scaling** goes back to Wegner (1972), who applied it in equilibrium critical phenomena. The Levenberg-Marquardt algorithm (Press *et al.* 2007) is probably the fitting routine most frequently employed for matching the estimates (with their error bars) from the simulation to the fitting function Eq. (7.14). There are a number of problems that can occur:

- Unless the result is purely qualitative, a good quality fit cannot be achieved without good quality numerical data, that includes a solid estimate of the numerical error, *i.e.* the estimated standard deviation of the estimate.
- The very setup of fitting function Eq. (7.14) (sometimes referred to as "the model") can introduce a systematic error; after all it is only a hypothesis.
- If $n > \tau - 1$ is very small, corrections due to the presence of the lower cutoff (s_0, Eq. (7.3)) can be very pronounced.
- The error stated for the fitted exponents alone can be misleading. If Eq. (7.14) is very constraining, the error will be low, but so will the goodness-of-fit.
- Too many fitting parameters allow for a very good goodness of fit, but also produce very large estimated statistical errors for the exponents.
- Fitting against a function with many parameters often is highly dependent on the initial guess. In order to achieve good convergence and systematic, controlled results, it may pay off to fit the data against Eq. (7.14) step-by-step, using the estimates obtained in a fit with fewer corrections as initial guesses for a fit with more corrections.
- In most cases, there is little point in having as many parameters as there are data points, as it often produces a seemingly perfect fit (goodness-of-fit of unity), independent of the input data.

[12] One might challenge the rule of thumb of the linear system size L having to span at least three orders of magnitude — in higher dimensions, say $d = 5$, spanning three orders of magnitude in linear extent leads to 15 orders of magnitude in volume, which might be the more suitable parameter to fit against.

- Extremely accurate data, *i.e.* estimates for the moments with very small error bars, may require a large number of correction terms.
- It can be difficult to force the corrections ω_i to be positive. It is not uncommon to fix them at certain reasonable values such as $\omega_i = i$ or $\omega_i = i/2$. Alternatively, they can be introduced differently, writing them, for example, in the form $\omega_i = i + |\tilde{\omega}_i|$.
- If finite size scaling applies, the relative statistical error for any moment scales like $\langle s^{2n} \rangle / \langle s^n \rangle^2 \propto L^{D(\tau-1)}$, assuming that $\sigma^2(s^n)$ scales like $\langle s^{2n} \rangle$, which it certainly does for $\tau > 1$. At $\tau = 1$ the scaling of $\sigma^2(s^n)$ may be slower than that of $\langle s^n \rangle^2$. While $L^{D(\tau-1)}$ does not depend on n, the amplitude of the moments does, leading normally to an increase of the relative error with n.

In some models the first moment of the avalanche size displays anti-correlations and thus faster numerical convergence as found in a mutually independent sample (Welinder *et al.* 2007). In many models, the average avalanche size $\langle s \rangle$ is known exactly, in one dimension often including the confluent singularities (Pruessner 2012a). These exact results can provide a test for convergence in numerics and also provide a **scaling relation**

$$D(2 - \tau) = \mu_1 \tag{7.15}$$

If μ_1 is known exactly ($\mu_1 = 2$ for bulk driving Manna, Oslo and Abelian BTW Models, $\mu_1 = 1$ for boundary drive), then Eq. (7.15) gives rise to a **scaling relation**. Normally, there are no further, strict scaling relations. However, the assumption of narrow joint distributions suggests $D(\tau - 1) = z(\alpha - 1)$ *etc.* (Christensen *et al.* 1991, Chessa *et al.* 1999a, Pruessner and Jensen 2004). If the exponent μ_1 is given by a mathematical identity and $\langle s \rangle$ serves as an analytically known reference in the numerical simulation, then μ_1 should not feature in the numerical analysis to extract the scaling exponents D and τ. Rather, when fitting μ_n versus $D(1 + n - \tau)$, this should be replaced by $D(n - 1) + \mu_1$.

Fitting μ_n in a linear fit (without corrections) against $D(1 + n - \tau)$ (or against $D(n - 1) + \mu_1$ if μ_1 is known exactly) is, in fact, the third step in the procedure described in this section. In principle, the $n > \tau - 1$ do not need to be integer valued. They have to be large enough to avoid a significant corrections due to the lower cutoff, and small enough to keep the relative statistical error small. Non-integer n can be computationally expensive, as they normally require at least one library call of `pow`.

While each estimate μ_n for every n should be based on the entire ensemble, considering them together in the same fit to extract the exponents D and τ introduces correlations, which are very often unaccounted for. As a result both goodness-of-fit as well as the statistical error for the exponents extracted are (unrealistically) small.

There are a number of strategies to address this problem. The simplest is to up-scale the error of the μ_n as if every estimate was based on a separate,

independent set of raw data. Considering M moments simultaneously, their error therefore has to scaled up by a factor $\sqrt{M}$ (Huynh *et al.* 2011). In a more sophisticated approach, one may extract estimates from a series of sub-samples (Efron 1982, Berg 1992, 2004).

It often pays to go through the process of extracting the exponents D and τ at an early stage of a simulation campaign, to identify potential problems in the data. Typical problems to watch out for include

- Corrections are too strong because system sizes are too small.
- Results are too noisy because sample sizes are too small, often because the system sizes are too big.
- Results have so little noise that fitting functions need to contain too many free parameters.
- Too few data points (*i.e.* too few different system sizes L or different moments n).
- Large event sizes suffer from integer overflow, resulting in seemingly negative or very small event sizes.
- Data identical in supposedly different runs, because of using the same seed for the random number generator.
- Transients chosen too short.

7.2.3 Statistical errors from chunks

One of the key-ingredients in the procedures described above is a reliable estimate for the statistical error of the estimates of the individual moments. The traditional approach is to estimate the variance, $\sigma^2(s^n) = \langle s^{2n} \rangle - \langle s^n \rangle^2$ of each moment, so that the statistical error of the estimate of $\langle s^n \rangle$ is estimated by $\sigma^2(s^n)/\sqrt{N/(2\tau+1)}$, where $N/(2\tau+1)$ is the number of effectively independent elements in the sample with correlation time τ.

This approach has a number of significant drawbacks. Firstly, each moment $\langle s^n \rangle$ requires a second moment, $\langle s^{2n} \rangle$, to be estimated as well. Considering a range of moments, this might (almost) double the computational effort. Rather dissatisfyingly, the highest moment estimated itself cannot be used to extract its finite size scaling exponent μ_n, because its variance is not estimated. Furthermore, because of their very high powers, the moments entering the estimates of the variances and thus the variances themselves have large statistical errors and are prone to integer overflow.

Estimating the effective number of independent elements in the sample is a hurdle that can be very difficult to overcome. Usually, it is based on an estimate of the correlation time τ. If $\langle s_i s_j \rangle - \langle s \rangle^2 = \sigma^2(s) \exp(-|i-j|/\tau)$, then the variance of the estimator

$$\overline{s} = \frac{1}{N} \sum_i^N s_i \tag{7.16}$$

of $\langle s \rangle$ for $N \gg \tau$ is

$$\sigma^2(\overline{s}) = \frac{1}{N^2} \sum_{ij}^{N} \left(\langle s_i s_j \rangle - \langle s \rangle^2 \right) \approx \frac{1 + \exp(-1/\tau)}{N(1 - \exp(-1/\tau))} \sigma^2(s) \approx \frac{2\tau + 1}{N} \sigma^2(s) \tag{7.17}$$

as if the sample contained only $N/(2\tau + 1)$ independent elements.

The main difficulty of this strategy is a reliable estimate of τ which often cannot be easily extracted from $\langle s_i s_j \rangle - \langle s \rangle^2$ because of noise and the presence of other exponential contributions, of which $\exp(-|i - j|/\tau)$ is the slowest decaying one. Moreover, in principle τ has to be measured for each observable separately (even when it makes physically most sense to assume that the system is characterised by a single correlation time).

To avoid these difficulties, one may resort to a simple sub-sampling plan. As discussed below (also Sec. 7.3.5), it is a matter of mere convenience and efficiency to repeatedly write estimates of moments based on a comparatively small sample into the output stream of a simulation and reset the cumulating variables. In the following these raw estimates based on a small sample are referred to as **chunks**. If their sample size is significantly larger than the correlation time, then each of these estimates can be considered as independent and the overall estimates based on it has its statistical error estimated accordingly. For example, if m_i with $i = 1, 2, \ldots, M$ are estimates of $\langle s^n \rangle$ all based on samples of the same size N, say $m_i = \sum_j^N s_{ij}^n$ with s_{ij} the jth element of the i sample, then the overall unbiased and consistent estimator (Brandt 1998) of $\langle s \rangle$ is

$$\overline{m} = \frac{1}{M} \sum_{i}^{M} m_i \tag{7.18}$$

which has an estimated standard deviation of $(\overline{m^2} - \overline{m}^2)/(M - 1)$ where

$$\overline{m^2} = \frac{1}{M} \sum_{i}^{M} m_i^2 \ . \tag{7.19}$$

One crucial assumption above is that the m_i are independent, which can always be achieved by merging samples. As long as M remains sufficiently large, one may be generous with the (effective) size of the individual samples (Flyvbjerg and Petersen 1989).

Chunks also allow a more flexible approach to determining and discarding transient behaviour from the sample supposedly taken in the stationary state. The transient can be determined as a (generous) multiple of the time after which (ideally all or several) observables no longer deviate more from the asymptotic or long time average than their characteristic variance. Where observables are known exactly (e.g. the average avalanche size Pruessner 2012a), they can be used as a suitable reference. Figure 7.5 shows the transient behaviour of the average avalanche size in a realisation of the Manna Model. A

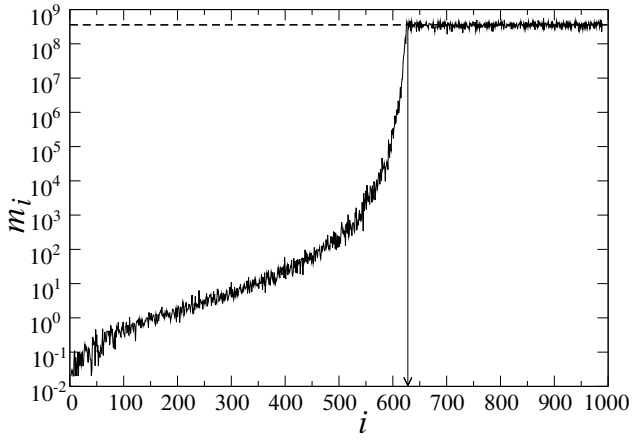


Fig. 7.5: Example of the transient behaviour of an observable (here the average avalanche size in the one-dimensional Manna Model with $L = 65536$) as a function of the chunk index in a log-lin plot (data from Huynh *et al.* 2011). The straight dashed line shows the exact expected average $\langle s \rangle$, Eq. (7.1). The arrow indicates the chunk from where on stationarity is roughly reached. A generous multiple of that time should be taken as the number of chunks to discard in order to ensure that correlations (and thus dependence on the initial setup) are essentially overcome.

more cautious strategy is to consider a series of different transients and study the change in the final estimates (with their estimated error) as a function of the transient discarded.

7.3 Algorithms and data organisation

In the following, a range of numerical and computational procedures are discussed that are commonly used in the numerical implementation of SOC models (for a more extensive review see Pruessner 2012b). Some of them are a matter of common sense and should be part of the coding repertoire of every computational physicist. However, it is not always entirely obvious how these "standard tricks" are used for SOC models.

In the following, the focus is on computational performance, which often comes with the price of lower maintainability of the code. The amount of real time spent on writing code and gained by making it efficient, should account for the time spent on debugging and maintaining it.

Most of the discussion below is limited to algorithmic improvements. The aim is produce code that communicates only minimally with the "outside

world", because in general, interaction with the operating system, as required for writing to a file, is computationally expensive and extremely slow. The UN*X operating system family (including, say, Linux and Mac OS X) distinguishes two different "modes" by which an executable keeps the CPU busy: By spending time on the (operating) system and by spending it in "user mode". Roughly speaking, the former accounts for any interaction between processes, with external controls or peripherals, including writing files. The latter accounts for the computation that takes place solely on the CPU (ALU, FPU, GPU, *etc.*) and the attached RAM. Tools like `time` and library functions like `getrusage` provide an interface to assess the amount of various resources used, while being themselves or resulting in systems calls.

Of course, the literature of computational physics in general is vast. Reviews and texts that are of particular use in the present context include Kernighan and Ritchie (1988), Cormen *et al.* (1996), Knuth (1997), Newman and Barkema (1999), Berg (2004), Landau and Binder (2005), Press *et al.* (2007).

7.3.1 Stacks

The definition of most SOC models makes no reference to the method to identify active sites, *i.e.* sites that are due to be updated. In principle, an implementation of an SOC model could therefore repeatedly scan the entire lattice to find the relevant sites. This is, however, very inefficient and therefore should be avoided. Instead, the index of active sites (or their coordinates) should be organised in a list. Every site in that list is subsequently updated. Moreover, it is often very important to know whether a site is maintained in the list or not. Sometimes this can be determined implicitly (for example, when a site is guaranteed to reside on the list from the moment its height exceeds the threshold), sometimes this is done explicitly by means of a flag associated with the site. The following contains a more detailed discussion of the various techniques available.

The most commonly used form of a list is a **stack**, called so, because this is how it appears to be organised. It consists of a vector, say `int stack[STACK_SIZE]`, of pre-defined size `STACK_SIZE`. It must be large enough to accommodate the maximum number of simultaneously active sites. Simulating large lattices, a balance has to be struck between what is theoretically possible and what is happening in practise.

The type of the stack, vector of `int` in the example above, is determined by the type it is meant to hold. If it holds the index of active sites, it is likely to be be `int`, but it may also hold more complex objects, say, coordinates of active particles (but see below). The number of objects currently held by the stack is stored in `int stack_height`.

If `STACK_SIZE` is smaller than the theoretical maximum of active sites, `int stack_height` has to be monitored as to prevent it from exceeding `STACK_SIZE`. The outcome of the simulation is undefined if that happens, because the exact position in memory of `stack[STACK_SIZE]` is *a priori* unknown. If therefore `stack_height` exceeds `STACK_SIZE`, memory has to be extended one way or another. For example, one may use `realloc()`, which assumes, however, that enough memory is actually available. Modern operating systems all provide virtual memory which is transparently supplemented by a swap file residing on the (comparatively slow) hard drive. This is to be avoided because of the computational costs associated. It may thus pay off for the process itself to make use, temporarily, of a file to store active sites. The alternative to abandon the particular realisation of the simulation introduces a bias away from rare events which is likely to have significant effect on observables. The same applies obviously if activity is suppressed if it reaches the maximum level.

There are two fundamental operations defined on a stack,

```
#define PUSH(a) stack[stack_height++]=(a)
#define POP(a) (a)=stack[--stack_height]
```

where `PUSH(a)` places `(a)` on the stack and `POP` takes an element off. The underlying idea is literally that of a stack: When a site becomes active, its index goes on a pile (`PUSH`) so that each index number on that pile represents a site waiting to be updated. When that happens, it is removed from the pile (`POP`).

It simplifies the code greatly if all objects on the stack are, in a sense, equivalent. For example, all sites on a stack are active. Guaranteeing this is not necessarily trivial, because the manipulation of one item on the stack may affect the state (and thus the eligibility) of another item on the stack. It is therefore advisable to ensure that all elements on the stack are distinct. In SOC models that means that active sites enter the stack exactly once, namely when the *turn* active. If an active site is charged again by a toppling neighbour, a new copy of its index is *not* placed on the stack. In the Manna Model, for instance, the single line of code to place objects on the stack could be

```
if (z[i]++==1) {PUSH(i);}
```

so that the index `i` of a site enters when it is charged while its height z is *at* the critical value z^c. The line should *not* read `if (z[i]++>=1)PUSH(i);`.

Unfortunately, the very data structure of a stack, which in the present context may better be called a **LIFO** (last in, first out), suggests a particular procedure to explore active sites, namely a depth first search (DFS); Whenever a toppling site activates its neighbours, one of them will be taken off first by the next call of `POP`, toppling in turn. Activity thus spreads very far very quickly, then returning, then spreading far again, rather than "burning locally". In fact, in DS-FFM a DFS is probably the simplest way of exploring a cluster of trees.

The alternative, a breadth first search (BFS) requires slightly greater computational effort because it normally makes use of a **FIFO** (first in, first out). The last object to arrive on a FIFO is the last one to be taken off, exactly the opposite order compared to a stack. Naively, this may be implemented by removing items from the front, `stack[0]`, and using `memmove()`[13] to feed it from the end, lowering `stack_height`. This approach, however, is computationally comparatively costly. A faster approach is to organise the stack in a queue, organised in a ring (circular buffer) to keep it finite, where a string of valid data grows at the end while retreating from the front.

In Abelian models, where the statistics of static features of avalanches, such as size and area, do not depend on the details of the microscopic dynamics[14], working through the stack using `POP` may be acceptable. Where temporal features are of interest too, the microscopic dynamics must implement a suitable microscopic time scale. Often the microscopic timescale is given by Poissonian updates, for example by active sites toppling with a Poissonian unit rate.

In principle that means that waiting times between events (sites toppling) are themselves random variables. If a faithful representation of the microscopic time is desired, then the random waiting times can be generated by taking the negative logarithm of a random number drawn from a uniform distribution on $(0, 1]$. If an approximate representation of the Poisson processes is acceptable (which, in fact converges to the exact behaviour in the limit of large numbers of active sites, see Liggett 2005), then elements are taken off the stack at random and time is made to progress in steps of `1./stack_height`. If `stack_height` remains roughly constant, than on average `stack_height` events occur per unit time as expected in a Poisson process. A simple implementation reads

```
int rs_pos;
#define RANDOM_POP(a) rs_pos=rand() % stack_height; (a)=stack[rs_pos
↳     ]; POP(stack[rs_pos])
```

where the last operation, `POP(stack[rs_pos])` overwrites the content of `stack[rs_pos]` by `stack[stack_height-1]` decrementing `stack_height` at the same time. When selecting the random position on the stack via `rs_pos=rand()% stack_height` a random number generator has to be used (Sec. 7.3.4), which only for illustrative purposes is called `rand()` here.

One consequence of the constraint of distinct objects on the stack is that a site may need to topple several times before being allowed to leave the stack. In Abelian models some authors circumvent that by placing a copy of the site index on the stack every time a pair of particles has to be toppled from it, which can be implemented easily by removing an appropriate number of

[13] Dedicated library functions like `memmove` and `memcpy` are generally much faster than naive procedures based on loops, although the latter can be subject to significant optimisation by the compiler.

[14] But note the strict definition of Abelianness discussed on p. 257.

particles from the site each time it enters the stack. As a result, however, stacks may become much larger, *i.e.* a greater amount of memory has to be allocated to accommodate them.

Depending on the details of the microscopic dynamics, an possible alternative is to relax a site completely after it has been taken off the stack, for example in the Manna Model:

```
while (stack_height) {
  RANDOM_POP(i);
  do {
    topple(i); /* Site i topples, removing two particles from i. */
    avalanche_size++; /* avalanche_size counts the number of topplings. */
  } while (z[i]>1);
}
```

where `topple(i)` reduces `z[i]` by 2 each time. If the avalanche size counts the number of topplings performed, `avalanche_size` has to be incremented within the loop. Counting only *complete* relaxations would spoil the correspondence with exact results.

An alternative approach with different microscopic time scale is to topple a site on the stack only once, and take it off only once it is fully relaxed. This approach requires some "tempering" with the stack:

```
while (stack_height) {
  i=rand() % stack_height;
  topple(stack[i]);
  if (z[i]<=1) POP(stack[i]);
}
```

In systems with parallel update, where all sites at the beginning of a time step have to be updated concurrently before updating the generation of sites that have been newly activated, a red-black approach (Dowd and Severance 1998) can be adopted. This requires the use of two stacks, which have to be swapped after completing one:

```
int *stack, stack_height=0;
int rb_stack[2][STACK_SIZE], next_stack_height;
int current_stack, next_stack;

#define NEXT_PUSH(a) rb_stack[next_stack][next_stack_height++]=(a)
#define NEXT_POP(a) (a)=rb_stack[next_stack][--next_stack_height]

...
current_stack=0;
next_stack=1;
stack=rb_stack[current_stack];
...
PUSH(i);
...
for (;;) {
  while (stack_height) {
    ...
    POP(i);
```

```
    ...
    NEXT_PUSH(j);
    ...
  }
  if (next_stack_height==0) break;
  /* Swap stacks. */
  stack_height=next_stack_height;
  next_stack_height=0;
  current_stack=next_stack;
  stack=rb_stack[current_stack];
  next_stack=1-next_stack;
}
/* Both stacks are empty. */
```

The use of the pointer `stack` is solely for being able to use the macros `PUSH` and `POP` defined earlier. Otherwise, it might be more suitable to define macros `CURRENT_PUSH` and `CURRENT_POP` corresponding to `NEXT_PUSH` and `NEXT_POP`.

A stack should also be used when determining the area of an avalanche, *i.e.* the number of distinct sites toppled (or visited, *i.e.* charged). To mark each site that has toppled during an avalanche and to avoid double counting, a flag has to be set, say `visited[i]=1` or `site[i].visited=1` (see Sec. 7.3.2). Counting how often the flag has been newly visited then gives the avalanche area. However, in preparation for the next avalanche, the flags have to be reset. This is when a stack comes handy, say

```
int area_stack[SYSTEM_SIZE];
int area_stack_height=0;
#define AREA_PUSH(a) area_stack[area_stack_height++]=(a)
#define AREA_POP(a) (a)=area_stack[--area_stack_height]
...
/* For each toppling site. */
if (visited[i]==0) {
  visited[i]=1;
  AREA_PUSH(i);
}
...
/* After the avalanche has terminated.
 * area_stack_height is the avalanche area. */
...
/* Re-initialise */
while (area_stack_height) {
  AREA_POP(i);
  visited[i]=0;
}
...
```

In the example above, the area is tracked implicitly in `area_stack_height`. The re-initialisation can be further improved using `while (area_stack_height )visited[area_stack[--area_stack_height]]=0`.

7.3.2 Sites and Neighbours

In SOC models, every site has a number of properties, most importantly the local degree of freedom, but also (statistical) observables which are being measured and updated as the simulation progresses. Other information associated with each site are flags (such as the one mentioned above to indicate whether a site had been visited) and even the neighbourhood (discussed below). In fact, the site itself may be seen as the **key** associated with all that information. That key might represent information in its own right, say, the coordinate, it might be an index of a vector, or a pointer.

7.3.2.1 Pointers and structures

A word of caution is in order with regard to pointers. The programming language C lends itself naturally to the use of pointers. However, code on the basis of pointers is difficult to optimise automatically at compile time. Depending on the quality of the compiler and the coding an index based implementation (which is also more portable) may thus results in faster code than the seemingly more sophisticated implementation based on pointers.

That said, in theory placing pointers on the stack, which gives immediately access to a relevant object should be faster than using indices, which are effectively an offset relative to a base: `b=z[stack[i]]` might result in machine code of the form `b=*(z+*(stack+i))` which contains one more addition than `b=*stack[i]` resulting in `b=**(stack+i)` if `stack` is a vector of pointers.

Similar considerations enter when using structures, which provide very convenient and efficient ways of organising and encapsulating data associated with each site. For example

```
struct site_struct {
  int height;
  char visited;
};
```

defines a structure with two members, `height` and `visited`. Declaring a variable `struct site_struct site[10]` allows the individual elements to be accessed in a structured way, say `site[i].height++, site[i].visited=1`. There are a number of computational drawbacks, which are, however, normally outweighed by the better maintainability of the code.

- Depending on the platform and the compiler, padding might become necessary, *i.e.* some empty space is added to the structure (Sec. 7.3.2.2, p. 292). The memory requirements of the structure is thus greater than the memory requirements for each variable when defined individually.
- Again depending on the platform as well as the compiler, without padding some operations on some types may require more CPU cycles (in particular when floating point types are used).

- Members within the structure are accessed similar to elements in a vector, namely by adding an offset. Access to the first member (where no offset is needed, `site[i].height` in the example above) can thus be faster than access to the other members (`site[i].visited` above). Because of that additional addition, the approach is often slower than using separate vectors for each member of the structure.

7.3.2.2 Neighbourhood information

It can be convenient, in particular for complicated topologies or when the neighbourhood information is computed or supplied externally, to store information about the local neighbourhood in a site structure, for example:

```
struct site_struct {
  ...
  int neighbour[MAX_NEIGHBOURS];
  int num_neighbours;
};
```

Because of the significant memory requirements, this is often not viable for large lattices. Again, instead of addressing neighbours by their index, pointers can be used, which often produces very efficient and elegant code.

The neighbours of each site thus are calculated and stored at the site only once. The strategy of pre-calculated neighbourhoods goes back to the very beginning of computational physics, when access to memory was much faster than doing such calculations on-the-fly.[15] This, however, has changed. It can be *much* faster to determine a neighbourhood on-the-fly than looking it up, unless, of course, the topology is so complicated that it becomes computationally too costly. Unfortunately, it is often difficult to try out different implementations (lookup tables and calculation on the fly), as the setup of a neighbourhood is at the heart of a lattice simulation.

As for calculating neighbourhoods, in one dimension the index of a site, which is strictly only a key to access the information, is often associated with its position on a one-dimensional lattice. Actual computation takes place only at boundaries. If the right neighbour of site `i` in the bulk is `i+1`, it may not exist on the right boundary or be `0` if periodic boundary conditions (PBC) apply in an implementation in C where the index of a vector of size `LENGTH` can take values from `0` to `LENGTH-1`. Similarly, the left neighbour is `i-1` in the bulk and `LENGTH-1` at `i=0` in case of periodic boundaries. Those are most easily implemented in the form `left=(i+LENGTH-1)%LENGTH` and `right=(i+1)%LENGTH` respectively using a modulo operation. The shift by `LENGTH` in the former avoids problems with negative indices at `i=0`.

A less elegant but often faster implementation is to determine whether a site is at the boundary before assigning the value for the neighbour, such as

[15] Back in the days when lookup tables for modulo operations were in fashion.

```
if (i==0) left=LENGTH−1;
else left=i−1;
```

or just `left=(i==0)?LENGTH-1:i-1`, which is more readable. This method is also more flexible with respect to the boundary condition implemented. Reflecting boundary conditions, for example are implemented by `left=(i==0)? 1 : i-1`. Open boundary conditions, on the other hand, might require special attention. If possible, they are best implemented using **padding**, *i.e.* by pretending that a neighbouring site exists, which, however, cannot interact with the rest of the lattice, for example, by making sure that it never fulfils the criterion to enter the stack. Such a site may need to be "flushed" occasionally to prevent it, for example, from fulfilling the criterion due to integer overflow. One might either assign one special site, say the variable `dump` in `left=(i==0)? dump : i-1` or allocate memory for `LENGTH+2` sites with an index from `0` to `LENGTH+1`, with valid sites ranging from `1` to `LENGTH` with sites `0` and `LENGTH+1` receiving charges without toppling in turn. This procedure also allows a very efficient way to determine the number of particles leaving the system, the **drop number** (Kadanoff *et al.* 1989).

Usually only in higher dimensions, one distinguishes reflecting boundary conditions, where the particle offloaded is moved to another site (normally the mirror image of the "missing" site), and "closed" boundary conditions, where the number of nearest neighbours is reduced and shed particles are evenly re-distributed among them.

Most of the above techniques remain valid in higher dimension, where the data can be organised in either a one-dimensional vector or a multidimensional vector. The former strategy makes use of macros of the form

```
#define COORDINATE2INDEX(x,y,z) ((x)+(LENGTH_X∗((y)+
↳     LENGTH_Y∗(z))))
#define INDEX2COORDINATE(i,x,y,z) z=(i)/(LENGTH_X∗LENGTH_Y),
↳     y=((i)/LENGTH_X)%LENGTH_Y,x=(i)%LENGTH_X
```

The use of the coma operator in the second macro helps to avoid errors when omitting curly brackets in expressions like `if (1)INDEX2COORDINATE(i,x,y,z);`. Where stacks are used to hold coordinates, the multiple assignments needed to store and fetch all of them may computationally outweigh the benefit of not having to calculate coordinates based on a single index.

The two biggest problem with the use of multi-dimensional vectors is their ambiguity when used with fewer indices and the consistency when passing them to functions. Both subtleties arise because of the logical difference between a vector of pointers to a type and the interpretation of a lower-dimensional variant of a multi-dimensional vector. While C makes that distinction, there is no syntactical difference between the two. For example

```
int a[2][10];

a[0][5]=7;
```

is a multi-dimensional vector using up `2*10*sizeof(int)` sequential bytes of memory. Each `a[i]` is the starting address of each row $i = 0, 1$. On the other hand

```
int *a[2];
int row1[10], row2[10];
a[0]=row1; a[1]=row2;

a[0][5]=7;
```

makes `a` a vector of pointers, using up `2*sizeof(* int)` bytes of memory, while each row uses `10*sizeof(int)` bytes. Both snippets of code declare `a` to be completely different objects, yet, for all intents and purposes in both cases `a` will behave like a two-dimensional array. That is, until it is to be passed as an argument to another function. In the first case, that function can be declared by `function(int array[2][10])`, informing it about the dimensions of the array, and subsequently called using `function(a)`. The two-dimensional vector `a` will behave as in the calling function. In fact, the function will even accept any other vector, lower dimensional or not, passed on to it as an argument (even when the compiler may complain).

In the second case, `a` is a vector of pointers to `int`, and so a function taking it as an argument must be declared in the form `function(int **a)`, using additional arguments or global constants (or variables) to inform it about the size of the vector. The two versions of the functions are incompatible, because a two-dimensional vector is really a one-dimensional vector with a particularly convenient way of addressing its components. In particular, the two-dimensional vector cannot be passed to the function designed for the second case using, say, `function(&a)` or `function((int **)a)`.

While these issues normally are resolved at the time of coding they can cause considerable problems when the memory allocation mechanism for the vector is changed. This happens, in particular, when lattice sizes are increased during the course of a simulation campaign. Initially, one might be tempted to define a lattice globally (stored in BSS or data segment) or as automatic variables taken from the stack, choosing a multi-dimensional array for convenience. Later on, they make be taken from the (usually much bigger) heap using `malloc()`, at which point the way they are accessed may have to be changed. The latter approach is the most flexible but possibly not the most convenient way of allocating memory for large items.

Finally, it is advisable to scan sites (when sweeping the lattice is unavoidable or scanning through a local neighbourhood) in a way that is local in memory and thus cache. The first option, declaring a two-dimensional vector in a single step, makes that more feasible than the second option, where different rows might end up at very different regions of memory. Not using higher dimensional vectors at all, however, is probably the best performing option.

7.3.3 Floating Point Precision

Very little and at times too little attention is being paid to the effect of limited floating point precision. Most SOC models can be implemented fully in integers even when their degrees of freedom are meant to be real valued, such as the Zhang Model (Zhang 1989), the Bak-Sneppen Model (Bak and Sneppen 1993) or the Olami-Feder-Christensen Model Model (Olami *et al.* 1992). In case of the latter, floating point precision has been found to significantly affect the results (Drossel 2002).

Where **random floating point** numbers are drawn, they might in fact contain much fewer random bits than suggested by the size of their mantissa. In that case, an implementation in integers is often not only faster but also "more honest". Where rescaling of variables cannot be avoided and occurs frequently, multiplying by a constant inverse often produces faster code than division.

Over the last decade or so, the floating point capabilities of most common CPUs have improved so much, however, that the difference in computational costs between integers and floating point arithmetics is either negligible or not clear-cut. The most significant disadvantage of the latter is the limited control of precision that is available on many platforms.

The levels of precision as defined in the IEEE standard 754 that are very widely used are single, double and extended. They refer to the number of bits in the mantissa determined when floating point operations are executed, *i.e.* they are the precision of the floating point unit (FPU). The precision the FPU is running at depends on platform, environment, compiler, compiler switches and the program itself. Some operating systems offer an IEEE interface, such as `fpsetprec()` on FreeBSD, and `fenv` on Linux.

Results of floating point arithmetics are stored in variables that may not offer the same level of precision the FPU is running at and in fact it is possible that none of the data types available matches a particular level of precision set on the FPU. Crucially, the precision setting of the FPU normally affects *all* floating point operations on *all* floating point variables, regardless of type, *e.g.* information is lost when results are calculated with extended precision and stored in variables offering only single precision. A notorious error observed on systems which default to extended precision, in particular Linux on x86, occurs when comparisons between variables produce different outcomes depending on the position in the code — at one point the result calculated may still reside on the FPU and thus offer extended precision, whereas at a later point the result is truncated after being written to memory. This can lead to serious inconsistencies when data is held in an ordered tree. Compiler switches like `-ffloat-store` for `gcc` help in these cases.

The commonly used `gcc` compiler offers three basic floating point types, `float`, `double` and `long double`, matching the three levels of precision mentioned above. The very nature of SOC means that observables span very many order of magnitudes. If variables that accumulate results, such as moments,

are too small (*i.e.* have a mantissa that is too small), smaller events may not accumulate at all any more once the variable has reached a sufficiently large value. This can skew estimates considerably where very large events occur very rarely. The macros `FLT_EPSILON`, `DBL_EPSILON` and `LDBL_EPSILON` in `float.h` give a suggestion of the relative scale of the problem. It can be mitigated by frequently "flushing" accumulating variables (see Sec. 7.3.5).

7.3.4 Random Number Generators

Random Number Generators (RNGs) are a key ingredient in many areas of computational physics, in particular in Monte-Carlo and Molecular Dynamics simulations. The vast majority of them, strictly, are not random, but follow instead a deterministic but convoluted computational path. RNGs are constantly being improved and evaluated, not least because of their use in cryptography. An introduction into the features of a good RNG can be found in the well-known Numerical Recipes (Press *et al.* 2007), with further details to be found in the review by Gentle (1998).

A "good" random number generator is one that offers a reasonable compromise between two opposing demands, namely that of speed and that of quality. In most stochastic SOC models, the RNG is used *very* often and thus typically consumes about half of the overall CPU time. Improving the RNG is thus a particularly simple way of improving the performance of an implementation. Because the variance (square of the standard deviation) of an estimate vanishes inversely proportional with the sample size it is based on, the performance of an implementation is best measured as the product of variance and CPU time spent "for it". However, one is ill-advised to cut corners by using a very fast RNG which has statistical flaws. The resulting problem may be very subtle and might not show until after a very detailed analysis.

One of the problems is the period of an RNG. Because RNGs generally have a finite state, they are bound to repeat a sequence of random numbers after a sufficient number of calls, at which point the simulation using the random numbers produces only copies of previous results. With improving hardware the RNG must therefore be re-assessed. A "good RNG" is a function of time, and very much a function of perception, as a mediocre RNG might appear to be a fantastic improvement over a poor RNG. It is good practise to use more than one random number generator to derive the same estimates and compare the results.

The C library's implementation of `rand()` is legendary for being unreliable and can be very poor. At the very least, it is essentially uncontrolled, although, of course, standards exist, which are, however, not always adhered to. It is fair to say that pure linear congruential RNGs are somewhat (out-)dated and indeed rarely used. They are, however, sometimes combined or

enhanced with more sophisticated techniques. In recent years, the Mersenne Twister (Matsumoto and Nishimura 1998, Matsumoto 2008) has become very widely used, yet, criticised by Marsaglia (2005) who proposed in turn KISS (Marsaglia 1999, but see Rose 2011), which is a remarkably simple RNG. The GNU Scientific Library (Galassi *et al.* 2009) contains an excellent collection of random number generators.

Somewhat more specific to the use of RNGs in SOC models is the frequent demand for random bits, for example in order to decide about the direction a particle is taking. Because every acceptable RNG is made up of equally random bits, each and everyone of them should be used for random booleans. These bits can be extracted one-by-one, by bit-shifting the random integer or by shifting a mask across, as in

```
#define RNG_MT_BITS (32)
#define RNG_TYPE unsigned long
RNG_TYPE mt_bool_rand=0UL;
RNG_TYPE mt_bool_mask=1UL<<(RNG_MT_BITS−1);
#define RNG_MT_BOOLEAN ( ( mt_bool_mask==(1UL<<(
↳     RNG_MT_BITS−1)) ) ? ((mt_bool_mask=1UL, mt_bool_rand=
↳     genrand_int32()) & mt_bool_mask) : (mt_bool_rand & (mt_bool_mask
↳     +=mt_bool_mask)) )
```

based on the Mersenne Twister. In general, bit shifts to the left using `a+=a` instead of `a<<=1` are faster, because the latter requires one more CPU cycle to write the constant `1` into the CPU's register.

More generally, integer random numbers have to be chosen uniformly from the range $\{0, 1, \ldots, n-1\}$ suggesting the use of the modulo operation, `r=rand()%n`. However, if `rand()` produces random integers uniformly from 0 up to and including `RAND_MAX`, then the modulo operation skews the frequencies with which random number occurs towards smaller values if `RAND_MAX+1` is not an integer multiple of n. The effect is of order $n/$`(RAND_MAX+1)` and thus is negligible if n is significantly smaller than `RAND_MAX`. However, picking a site at random on a very large lattice or an element from a very large stack, this effects becomes a realistic concern. In that case, the modulo operation can be used on a random number drawn uniformly among integers from 0 up to and including $R-1$, where R is a multiple of n and ideally the largest multiple of n less or equal to `RAND_MAX+1`:

```
const long long int n=...;
/* The constant multiple_minus_1 is made to have type as the return
 * value of rand(). */
const int multiple_minus_1=(n*((((long long int)RAND_MAX) + 1LL)/n)
↳     )−1LL;
int r;
#define RANDOM(a) while ((r=rand())>multiple_minus_1); (a)=r%n
```

where `multiple_minus_1` plays the rôle of $R-1$. When determining the maximum multiple, it is crucial that the operation `RAND_MAX+1` is performed using a type where the addition does not lead to rounding or integer overflow. The latter is also the reason why one is subtracted in the expression for

`multiple_minus_1`, which otherwise might not be representable in the same type as the return value of `rand()`, which is necessary to avoid any unwanted type casting at run time.[16]

The initial seed of the RNG needs to be part of the output of the programme it is used in, so that the precise sequence of events can be reproduced in case an error occurs. Some authors suggest that the initial seed itself should be random, based, for example, on `/dev/random`, or the library functions `time ()` or `clock()`,[17] and that the RNG carries out a "warm-up-cycle" of a few million calls (Jones 2012). After that, it is sometimes argued, chances are that one sequence of (pseudo) random numbers is independent from another sequence of random numbers generated by the same RNG based on a different seed. Fortunately, some RNGs, in particular those designed for use on parallel machines, offer a facility to generate sequences that are guaranteed to be independent. Where poor-man's parallel computing (many instances of the same simulation running with different seeds) takes place, independent sequences are of much greater concern than in situations where different parameter settings are used in different instances. In the former case the data of all instances will be processed as a whole, probably under the assumption that it is actually independent. In the latter case, the results will enter differently and using even an identical sequence of random numbers will probably not have a noticeable effect. All these caveats are put in perspective by the fact that most SOC models fed by a slightly differing sequences of pseudo random numbers take "very different turns in phase space" and thus will display very little correlations.

7.3.5 Output

As mentioned above, it is generally advisable to output and flush data frequently in **chunks**, resetting accumulating variables afterwards. Even when output occurs every second, the overhead in terms of the CPU and real time spent by the system is likely to be negligibly small.

Where data is written to a file in large quantities or frequently, **buffered I/O** as provided by `stdio` through the `printf`-family of library calls is usually much faster than writing immediately to the file using `unistd`'s `write`. There are two caveats to this approach: Firstly, depending on the size of the buffer and thus the frequency of writing, a significant amount of CPU time may be

[16] This is one of the many good reasons to use constants rather than macros (van der Linden 1994, Kernighan and Pike 2002).

[17] Both functions are bad choices on clusters where several instances of the same programme are intended to run in parallel. The function `time()` changes too slowly (returning the UN*X epoch time in seconds) and the function `clock()` wraps after about 36 minutes, so that neither function guarantees unique seeds. In general, seeding is best done explicitly.

lost if the program terminates unexpectedly. To avoid corrupt data, `fflush()` should be used rather than allowing the buffer to empty whenever it reaches its high-water mark. Secondly, if buffering I/O has a significant impact on the computational performance, the data may better be processed on-the-fly rather than storing it in a file.

In the following, `stdio` is used for its convenient formatting capabilities, provided by the plethora of flags in the formatting string of a `printf` call. To avoid the problems mentioned above, buffers are either flushed after each chunk by means of `fflush`, or buffering is switched to buffering line by line, using `setlinebuf`.

To avoid unexpected interference of the operating system with the simulation, operations should be avoided that can potentially fail because the environment changes. This applies, in particular, to read and write access to files. In any case, such operations need to be encapsulated in an `if` condition that catches failing system calls and triggers a suitable remedy.

Output of chunks should therefore happen through the `stdout` stream which is by default open at the time of the program start. As the output is usually used in post-processing it needs to be retained, which can be achieved by re-directing `stdout` into a file. In the typical shell syntax this can be done in the command line by, say, `./simulation > output.txt`. To allow easy post-processing, every line should contain all relevant simulation parameters, such as the system size, the number of the chunk (a counter), the number of events per chunk, the initial seed of the random number generator (RNG), in fact, everything that is needed to reproduce that line from scratch or to plot the relevant (derived) data. Typical examples are moments to be plotted against the system size and moment ratios, involving different moments of the same observable. Using post-processing tools to wade through vast amounts of data to find the missing piece of information to amend a line of data can require significant effort and is highly error-prone.

Repeating the same output (system size, RNG seed etc) over and over seemingly goes against the ethos of avoiding redundant information, which should be applied when setting up a computer simulation (to avoid clashes), but is wholly misplaced when it comes to data output. In fact, redundancy in output is a means to measure consistency and a matter of practicality as almost all basic post-processing tools are line-oriented.

In some rare cases, an action by the simulation or an event on the system can result in a **signal** being sent to the running instance of the program. In response the program suspends the current operation, executes a signal handler and continues where it left off. In principle, the signal should not lead to inconsistent data or behaviour; in fact, it is probably the most basic but also a very convenient way to communicate with a running program. For example

```
#include <signal.h>
...
void sighup_handler(int signo);
```

```
...
signal(SIGHUP, sighup_handler);
...
void sighup_handler(int signo)
{
finish_asap=1;
}
```

assigns the signal handler `sighup_handler` to deal with the signal `SIGHUP`, which can be sent to the program using `kill -HUP`.

There is a rare situation when the signal interrupts in a way that it leads to unexpected behaviour, namely when it arrives while a "slow system call" is executed, *i.e.* an operation that is performed by the kernel on behalf of the programme, but which can take a long time to complete, such as `pause` , `sleep`, but also `write` to so-called pipes. Without discussing the technical details of the latter, it can lead to inconsistencies in the output which might not be detected in the post-processing. For example, a chunk may contain truncated lines and thus may lack certain information or data, which the post-processing tools might treat as zeroes. Apart from a graphical inspection of the data, two measures may therefore be advisable: Firstly, output can be encapsulated in calls of `sigprocmask` which allows temporary suspension of the delivery of signals. Secondly, a chunk can be terminated by a single line containing a keyword to indicate the successful completion of the output (*i.e.* without catching an error, in particular not an "interrupted system call", `EINTR`), such as the tag (see below) `#Completed`. Simply counting the number of occurrences of that tag and comparing to (supposed) the number of valid chunks can pick up inconsistencies. In large scale simulations, where disk space can be a problem leading to truncated files as the system runs out of file space, this is particularly advisable.

After a chunk has been written out, variables collecting data have to be reset. Where PDFs are estimated, sweeping across the entire histogram can become expensive and therefore performing all relevant steps simultaneously is advantageous for the overall performance. Using one of the examples above (Sec. 7.2.1.1):

```
long long total=0;
for (i=0; i<SMALL2MEDIUM_THRESHOLD; i++)
  if (histo_small[i]) {
    printf(...);
    total+=histo_small[i];
    histo_small[i]=0;
  }
...
printf("out_of_range:␣%i\n", histo_out_of_range);
total+=histo_out_of_range;
histo_out_of_range=0;
printf("total:␣%lli", total);
```

The final line allows the user to compare the number of events collected in the histogram to the number of events expected. It is a computationally cheap additional check for data consistency.

To distinguish different types of output, such as moments of different observables, data should be **tagged** by short keys that are easily filtered out in post processing. For example, if every line containing moments of avalanche sizes is tagged by `#M_SIZE` at the beginning, all relevant lines can be extracted very easily for example using `grep '^#M_SIZE'output.txt`. To strip off the tags, one either appends `|sed 's/#M_SIZE//'` or includes the functionality of `grep` in the `sed` command,

```
sed -n 's/^#M_SIZE//p' output.txt > output.txt_M_SIZE
```

storing all relevant lines in `output.txt_M_SIZE` for further processing by other tools. One very simple, but particularly powerful one is `awk`. For example, the average across the seventh column starting with the 101st chunk (stored in the first column) can be calculated using

```
awk ' { if ($1>100) {m0++; m1+=$7;} } END { printf ("%i %10.20g\n",
↳     m0, m1/m0); } ' output.txt_M_SIZE
```

All of this is very easily automated using powerful scripting languages (in particular shell scripts, `awk`, `sed` and `grep`), and more powerful (interpreted) programming languages, such as `perl` or `python`, which provide easy access to line-oriented data. In recent years, XML has become more popular to store simulation parameters as well as simulation results.

7.4 Summary and conclusion

The early life of SOC was all about computer models that showed the desired features of SOC: Intermittent behaviour (slow drive, fast relaxation) displaying scale invariance as observed in traditional critical phenomena without the need to tune a control parameter to a critical value. After many authors had (mostly with little success) attempted to populate the universality class of the BTW Sandpile, a range of SOC models was proposed firstly as a paradigm of alternative universality classes and later to highlight specific aspects of SOC, such as non-conservation (as for example in the Forest-Fire Model), non-Abelianness (as for example in the Olami-Feder-Christensen Model) and stochasticity (as for example in the Manna Model).

Many of these models have been studied extensively, accumulating hundreds of thousands of hours of CPU time in large-scale Monte Carlo simulations. A finite size scaling analysis of the data generally produces a set of two to eight exponents, which are supposedly universal. It turns out, however, that very few models display clean, robust scaling behaviour in the event size distribution, although it is remarkably broad for many models.

Of the models discussed above, the Manna Model displays the clearest signs of scale invariance. There is wide consensus that it is the same universality class as the Oslo Model (Christensen *et al.* 1996, Nakanishi and Sneppen 1997). In the conservative limit and in the near-conservative regime, the Olami-Feder-Christensen Model also displays convincing moment scaling, but less so for smaller values of the level of conservation. Numerical artefacts may play a significant rôle in its scaling (Drossel 2002).

The Forest Fire Models is widely acknowledged for failing to display finite size scaling in the event size distribution (Grassberger 2002, Pruessner and Jensen 2002a), although its moments still display some scaling (Pruessner and Jensen 2004). The contrast is even sharper in the Bak-Tang-Wiesenfeld Model: Some scaling is known analytically (Majumdar and Dhar 1992, Ivashkevich 1994, Ivashkevich *et al.* 1994, Dhar and Manna 1994), yet the event size distribution seems at best be governed by multiscaling (Tebaldi *et al.* 1999, Drossel 1999, 2000, Dorn *et al.* 2001)

While analytical approaches receive increasing attention, numerical techniques remain indispensable in the development and analysis of models which are tailor-made to display specific features or to mimic experimental systems. Models developed more recently are usually implemented in C, producing numerical data in Monte-Carlo simulations. It is fair to say that the careful data analysis requires as much attention to detail as the implementation of the model in the first place.

While the classic data-collapse and more immediate tests for scaling dominated the early literature of SOC, more recently the finite size scaling of moments (Tebaldi *et al.* 1999) has become the predominant technique for the extraction of scaling exponents. Apart from identifying the mechanism of SOC, the main purpose of the numerical work is to establish universality and universality classes among models, as well as their relation to natural phenomena. One may hope that these efforts will eventually help to uncover the necessary and sufficient conditions for SOC.

Acknowledgements The author gratefully acknowledges the kind support by EPSRC Mathematics Platform grant EP/I019111/1.

Appendix: Implementation details for binning

To implement binning in computer simulations of SOC models it is advisable to perform simple bit manipulations on basic, integer-valued observables. It often suffices to implemented three levels of coarse graining or less, for example

```
#define SMALL2MEDIUM_THRESHOLD (1LL<<15)
long long histo_small[SMALL2MEDIUM_THRESHOLD]={0LL};
#define MEDIUM2LARGE_THRESHOLD (1LL<<30)
```

```
#define MEDIUM_SHIFT (12)
long histo_medium[(MEDIUM2LARGE_THRESHOLD−
↳     SMALL2MEDIUM_THRESHOLD)>>MEDIUM_SHIFT]={0L};
#define LARGE_THRESHOLD (1LL<<45)
#define LARGE_SHIFT (27)
int histo_large[(LARGE_THRESHOLD−MEDIUM2LARGE_THRESHOLD)
↳     >>LARGE_SHIFT]={0};
int histo_out_of_range=0;
long long int s; /* event size */

...

if (s<SMALL2MEDIUM_THRESHOLD) histo_small[s]++;
else if (s<MEDIUM2LARGE_THRESHOLD) histo_medium[(s−
↳     SMALL2MEDIUM_THRESHOLD)>>MEDIUM_SHIFT]++;
else if (s<LARGE_THRESHOLD) histo_large[(s−
↳     MEDIUM2LARGE_THRESHOLD)>>LARGE_SHIFT]++;
else histo_out_of_range++;
```

Here the event size to be tallied is `s`. In the block of `if` statements, it is compared to various thresholds before it is rescaled and counted into a histogram. Because vectors in many programming languages start with index 0, a shift an offset is subtracted as well. It can pay of to re-arrange the `if` statements as to test against the most frequent case as early as possible. One case, in the present example the last one, counts the number of times the counter overspills, here `histo_out_of_range`.

Some subtleties of the above implementation are worth discussing. Firstly, the types used for the histogram typically decrease in size with increasing event size while the size of the type needed to represent the event size at the respective thresholds increases. This is because normally the frequency is an *inverse* power law of the event size. Great care must be taken to avoid unnecessary typecasts and undesired outcomes, as some languages, in particular C, are rather idiosyncratic when it comes to (integer) type-promotion in comparisons, in particular when they involve signs.

In the above examples, automatic vector variables are used and initialised by assigning `{0}`, which is expanded by the compiler to a suitable size by adding zeroes. Initialisation of vectors in C has been further simplified in the C99 standard.

Secondly, it is important to choose the thresholds together with the planned bit-shifts, in order to avoid an *off-by-one* error. The problem is that, say,
`s<MEDIUM2LARGE_THRESHOLD`, does not imply

```
(s−SMALL2MEDIUM_THRESHOLD)/((1<<MEDIUM_SHIFT) < (
↳     MEDIUM2LARGE_THRESHOLD−SMALL2MEDIUM_THRESHOLD)
↳     /(1<<MEDIUM_SHIFT)
```

because for some `s<MEDIUM2LARGE_THRESHOLD` their bitshifted value `s>>MEDIUM_SHIFT` in fact equals `MEDIUM2LARGE_THRESHOLD>>MEDIUM_SHIFT`, namely precisely when `MEDIUM2LARGE_THRESHOLD` is not an integer multiple of

`1<<MEDIUM_SHIFT`. It is therefore a matter of defencive programming to write the thresholds for the macros in this form:

```
#define MEDIUM2LARGE_THRESHOLD ((1LL<<18) * (1LL<<
↳    MEDIUM_SHIFT))
```

As for a rudimentary output routine

```
for (i=0; i<SMALL2MEDIUM_THRESHOLD; i++)
  if (histo_small[i]) printf("%i %i %lli %i\n", i, i, histo_small[i], 1);
for (i=0; i<((MEDIUM2LARGE_THRESHOLD−
↳    SMALL2MEDIUM_THRESHOLD)>>MEDIUM_SHIFT); i++)
  if (histo_medium[i]) printf("%li %i %li %i\n", ((long)
↳    SMALL2MEDIUM_THRESHOLD)+(((long)(i))<<MEDIUM_SHIFT),
↳     i, histo_medium[i], 1<<MEDIUM_SHIFT);
for (i=0; i<((LARGE_THRESHOLD−MEDIUM2LARGE_THRESHOLD)
↳    >>LARGE_SHIFT); i++)
  if (histo_large[i]) printf("%lli %i %i %i\n", ((long long)
↳    MEDIUM2LARGE_THRESHOLD)+(((long long)(i))<<
↳    LARGE_SHIFT), i, histo_large[i], 1<<LARGE_SHIFT);
printf("out_of_range: %i\n", histo_out_of_range);
```

care must again be taken that the formatting of the output is in line with the type of the data and does not spoil it. Fortunately, most modern compilers spot clashes between the formatting string used in `printf` and the actual argument. As discussed below, it is generally advisable to have only one output stream, namely `stdout`, and to use **tags** to mark up data for easy fetching by post-processing tools. In the example above, the bins have not been rescaled by their size which instead has been included explicitly in the output. A sample of the PDF can be inspected by plotting the third column divided by the fourth against the first.

References

Alava, M., and M. A. Muñoz, 2002, Phys. Rev. E **65**(2), 026145 (pages 8).

Alava, M. J., L. Laurson, A. Vespignani, and S. Zapperi, 2008, Phys. Rev. E **77**(4), 048101 (pages 2), comment on (Pruessner and Peters 2006), reply (Pruessner and Peters 2008).

Bagnoli, F., F. Cecconi, A. Flammini, and A. Vespignani, 2003, Europhys. Lett. **63**(4), 512.

Bak, P., K. Chen, and M. Creutz, 1989a, Nature **342**(6251), 780.

Bak, P., K. Chen, and C. Tang, 1990, Phys. Lett. A **147**(5–6), 297.

Bak, P., K. Christensen, L. Danon, and T. Scanlon, 2002, Phys. Rev. Lett. **88**(17), 178501 (pages 4).

Bak, P., and K. Sneppen, 1993, Phys. Rev. Lett. **71**(24), 4083.

Bak, P., and C. Tang, 1989, J. Geophys. Res. **94**(B11), 15635.

Bak, P., C. Tang, and K. Wiesenfeld, 1987, Phys. Rev. Lett. **59**(4), 381.

Bak, P., C. Tang, and K. Wiesenfeld, 1989b, in *Cooperative Dynamics in Complex Physical Systems, Proceedings of the Second Yukawa International Symposium, Kyoto, Japan, August 24–27, 1988*, edited by H. Takayama (Springer-Verlag, Berlin, Germany), volume 43 of *Springer Series in Synergetics*, pp. 274–279.

Basu, M., U. Basu, S. Bondyopadhyay, P. K. Mohanty, and H. Hinrichsen, 2012, Phys. Rev. Lett. **109**, 015702.
Ben-Hur, A., and O. Biham, 1996, Phys. Rev. E **53**(2), R1317.
Berg, B. A., 1992, Comp. Phys. Comm. **69**(1), 7.
Berg, B. A., 2004, *Markov Chain Monte Carlo Simulations and Their Statistical Analysis* (World Scientific, Singapore).
Biham, O., E. Milshtein, and O. Malcai, 2001, Phys. Rev. E **63**(6), 061309 (pages 8).
Binder, K., and D. W. Heermann, 1997, *Monte Carlo Simulation in Statistical Physics* (Springer-Verlag, Berlin, Germany), 3rd edition.
Bonachela, J. A., 2008, *Universality in Self-Organized Criticality*, Ph.D. thesis, Departmento de Electromagnetismo y Fisíca de la Materia & Institute Carlos I for Theoretical and Computational Physics, University of Granada, Granada, Spain, accessed 12 Sep 2009, URL `http://hera.ugr.es/tesisugr/17706312.pdf`.
Bonachela, J. A., and M. A. Muñoz, 2007, Physica A **384**(1), 89, ISSN 0378-4371, proceedings of the International Conference on Statistical Physics, Raichak and Kolkata, India, Jan 5–9, 2007.
Bonachela, J. A., and M. A. Muñoz, 2009, J. Stat. Mech. **2009**(09), P09009 (pages 37).
Brandt, S., 1998, *Data Analysis* (Springer-Verlag, Berlin, Germany).
Bröker, H.-M., and P. Grassberger, 1999, Physica A **267**(3–4), 453, ISSN 0378-4371.
Bruinsma, R., and G. Aeppli, 1984, Phys. Rev. Lett. **52**(17), 1547.
Bunzarova, N. Z., 2010, Phys. Rev. E **82**(3), 031116 (pages 14).
Burridge, R., and L. Knopoff, 1967, Bull. Seismol. Soc. Am. **57**(3), 341.
Carlson, J. M., J. S. Langer, and B. E. Shaw, 1994, Rev. Mod. Phys. **66**(2), 657.
Casartelli, M., L. Dall'Asta, A. Vezzani, and P. Vivo, 2006, Eur. Phys. J. B **52**(1), 91.
Černák, J., 2002, Phys. Rev. E **65**(4), 046141 (pages 6).
Ceva, H., 1995, Phys. Rev. E **52**(1), 154.
Ceva, H., 1998, Phys. Lett. A **245**(5), 413.
Chessa, A., H. E. Stanley, A. Vespignani, and S. Zapperi, 1999a, Phys. Rev. E **59**(1), R12, numerics may not be independent from (Chessa *et al.* 1999b).
Chessa, A., A. Vespignani, and S. Zapperi, 1999b, Comp. Phys. Comm. **121–122**, 299, proceedings of the Europhysics Conference on Computational Physics CCP 1998, Granada, Spain, Sep 2–5, 1998. Numerics presented may not be independent from that in (Chessa *et al.* 1999a).
Chianca, C. V., J. S. Sá Martins, and P. M. C. de Oliveira, 2009, Eur. Phys. J. B **68**(4), 549.
Christensen, K., 1992, *Self-Organization in Models of Sandpiles, Earthquakes, and Flashing Fireflies*, Ph.D. thesis, Institute of Physics and Astronomy, University of Aarhus, DK-8000 Aarhus C, Denmark.
Christensen, K., 1993, Phys. Rev. Lett. **71**(8), 1289, reply to comment (Klein and Rundle 1993).
Christensen, K., Á. Corral, V. Frette, J. Feder, and T. Jøssang, 1996, Phys. Rev. Lett. **77**(1), 107.
Christensen, K., N. Farid, G. Pruessner, and M. Stapleton, 2008, Eur. Phys. J. B **62**(3), 331.
Christensen, K., H. C. Fogedby, and H. J. Jensen, 1991, J. Stat. Phys. **63**(3/4), 653.
Christensen, K., and N. R. Moloney, 2005, *Complexity and Criticality* (Imperial College Press, London, UK).
Christensen, K., and Z. Olami, 1992, Phys. Rev. A **46**(4), 1829.
Christensen, K., and Z. Olami, 1993, Phys. Rev. E **48**(5), 3361.
Clar, S., B. Drossel, and F. Schwabl, 1996, J. Phys.: Condens. Matter **8**(37), 6803.
Clauset, A., C. R. Shalizi, and M. E. J. Newman, 2009, SIAM Rev. **51**(4), 661.
Cormen, T. H., C. E. Leiserson, and R. L. Rivest, 1996, *Introduction to Algorithms* (The MIT Press, Cambridge, MA, USA).

Corral, Á., 2003, Phys. Rev. E **68**(3), 035102(R) (pages 4).
Corral, Á., 2004a, Phys. Rev. E **69**(2), 026107 (pages 12).
Corral, Á., 2004b, Phys. Rev. Lett. **92**(10), 108501 (pages 4).
Corral, Á., 2004c, Physica A **340**(4), 590.
Corral, Á., and K. Christensen, 2006, Phys. Rev. Lett. **96**(10), 109801 (pages 1), comment on (Lindman *et al.* 2005), reply (Lindman *et al.* 2006).
Corral, Á., and M. Paczuski, 1999, Phys. Rev. Lett. **83**(3), 572.
Creutz, M., 2004, Physica A **340**(4), 521 , ISSN 0378-4371, proceedings of the symposium Complexity and Criticality: in memory of Per Bak (1947–2002), Copenhagen, Denmark, Aug 21–23, 2003.
Davidsen, J., and M. Paczuski, 2005, Phys. Rev. Lett. **94**(4), 048501 (pages 4), comment (Werner and Sornette 2007).
Davidsen, J., and M. Paczuski, 2007, Phys. Rev. Lett. **99**(17), 179802, reply to comment (Werner and Sornette 2007).
De Menech, M., and A. L. Stella, 2000, Phys. Rev. E **62**(4), R4528.
De Menech, M., A. L. Stella, and C. Tebaldi, 1998, Phys. Rev. E **58**(3), R2677.
de Sousa Vieira, M., 1992, Phys. Rev. A **46**(10), 6288.
Dhar, D., 1990, Phys. Rev. Lett. **64**(14), 1613.
Dhar, D., 1999a, Physica A **263**(1–4), 4, proceedings of the 20th IUPAP International Conference on Statistical Physics, Paris, France, Jul 20–24, 1998, overlaps with (Dhar 1999b).
Dhar, D., 1999b, Studying self-organized criticality with exactly solved models, arXiv:cond-mat/9909009, `arXiv:cond-mat/9909009`.
Dhar, D., 2004, Physica A **340**(4), 535.
Dhar, D., 2006, Physica A **369**(1), 29, proceedings of the 11th International Summerschool on 'Fundamental Problems in Statistical Physics', Leuven, Belgium, Sep 4 – 17, 2005; updated from (Dhar 1999b).
Dhar, D., and S. S. Manna, 1994, Phys. Rev. E **49**(4), 2684.
Dhar, D., and P. Pradhan, 2004, J. Stat. Mech. **2004**(05), P05002 (pages 12), includes erratum.
Dhar, D., and R. Ramaswamy, 1989, Phys. Rev. Lett. **63**(16), 1659.
Dhar, D., P. Ruelle, S. Sen, and D.-N. Verma, 1995, J. Phys. A: Math. Gen. **28**(4), 805.
Dickman, R., and J. M. M. Campelo, 2003, Phys. Rev. E **67**(6), 066111 (pages 5).
Dickman, R., M. A. Muñoz, A. Vespignani, and S. Zapperi, 2000, Braz. J. Phys. **30**(1), 27.
Dickman, R., T. Tomé, and M. J. de Oliveira, 2002, Phys. Rev. E **66**(1), 016111 (pages 8).
Dickman, R., A. Vespignani, and S. Zapperi, 1998, Phys. Rev. E **57**(5), 5095.
Dorn, P. L., D. S. Hughes, and K. Christensen, 2001, On the avalanche size distribution in the btw model, preprint from `http://www.cmth.ph.ic.ac.uk/kim/papers/preprints/preprint_btw.pdf`, accessed 19 Oct 2010.
Dowd, K., and C. Severance, 1998, *High Performance Computing* (O'Reilly, Sebastopol, CA, USA), 2nd edition.
Drossel, B., 1999, An alternative view of the Abelian sandpile model, `arXiv:cond-mat/9904075v1`.
Drossel, B., 2000, Phys. Rev. E **61**(3), R2168.
Drossel, B., 2002, Phys. Rev. Lett. **89**(23), 238701 (pages 4).
Drossel, B., and F. Schwabl, 1992a, Phys. Rev. Lett. **69**(11), 1629, largely identical to proceedings article (Drossel and Schwabl 1992b).
Drossel, B., and F. Schwabl, 1992b, Physica A **191**(1–4), 47, proceedings of the International Conference on Fractals and Disordered Systems, Hamburg, Germany, Jul 29–31, 1992.

Efron, B., 1982, *The Jackknife, the Bootstrap and Other Resampling Plans* (SIAM, Philadelphia, PA, USA).
Feller, W., 1966, *An Introduction to Probability Theory and its Applications*, volume II (John Wiley & Sons, New York, NY, USA).
Flyvbjerg, H., and H. G. Petersen, 1989, J. Chem. Phys. **91**(1), 461.
Galassi, M., J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, and F. Rossi, 2009, *GNU Scientific Library Reference Manual* (Network Theory Ltd.), 3rd (v1.12) edition, `http://www.network-theory.co.uk/gsl/manual/`, accessed 18 Aug 2009.
Gentle, J. E., 1998, *Random Number Generation and Monte Carlo Methods* (Springer-Verlag, Berlin, Germany).
Grassberger, P., 1982, Z. Phys. B **47**, 365.
Grassberger, P., 1994, Phys. Rev. E **49**(3), 2436.
Grassberger, P., 1995, Phys. Lett. A **200**(3–4), 277.
Grassberger, P., 2002, New J. Phys. **4**(1), 17 (pages 15).
Grassberger, P., and H. Kantz, 1991, J. Stat. Phys. **63**(3–4), 685.
Grinstein, G., D.-H. Lee, and S. Sachdev, 1990, Phys. Rev. Lett. **64**(16), 1927.
Gutenberg, B., and C. F. Richter, 1954, *Seismicity of the earth and associated phenomena* (Princeton University Press, Princeton, NJ, USA), 2nd edition.
Henkel, M., H. Hinrichsen, and S. Lübeck, 2008, *Non-Equilibrium Phase Transitions* (Springer-Verlag, Berlin, Germany).
Henley, C. L., 1989, Bull. Am. Phys. Soc. **34**(3), 838, abstract of talk M18.2, 23 March 1989, of the 1989 March Meeting of The American Physical Society, St. Louis, MO, USA, Mar 20–24, 1989.
Henley, C. L., 1993, Phys. Rev. Lett. **71**(17), 2741.
Hergarten, S., 2002, *Self-Organized Criticality in Earth Systems* (Springer-Verlag, Berlin, Germany).
Hinrichsen, H., 2000, Adv. Phys. **49**, 815.
Honecker, A., and I. Peschel, 1997, Physica A **239**(4), 509.
Hughes, D., and M. Paczuski, 2002, Phys. Rev. Lett. **88**(5), 054302 (pages 4).
Huynh, H. N., and G. Pruessner, 2012, Phys. Rev. E **85**, 061133.
Huynh, H. N., G. Pruessner, and L. Y. Chew, 2011, J. Stat. Mech. **2011**(09), P09024.
Hwa, T., and M. Kardar, 1989a, Phys. Rev. Lett. **62**(16), 1813, identical to proceedings article (Hwa and Kardar 1989b).
Hwa, T., and M. Kardar, 1989b, Physica D **38**(1–3), 198, identical to (Hwa and Kardar 1989a); Proceedings of a conference held in honour of Benoit B. Mandelbrot's 65th birthday, Les Mas d'Artigny (Vence), France, Oct 1 – 4, 1989.
Hwa, T., and M. Kardar, 1992, Phys. Rev. A **45**(10), 7002.
Ito, K., and M. Matsuzaki, 1990, J. Geophys. Res. **95**(B5), 6853.
Ivashkevich, E. V., 1994, J. Phys. A: Math. Gen. **27**(11), 3643.
Ivashkevich, E. V., D. V. Ktitarev, and V. B. Priezzhev, 1994, Physica A **209**(3–4), 347.
Janssen, H. K., 1981, Z. Phys. B **42**, 151.
Jeng, M., 2005, Phys. Rev. E **71**(3), 036153 (pages 17).
Jensen, H. J., 1998, *Self-Organized Criticality* (Cambridge University Press, New York, NY, USA).
Jo, H.-H., and M. Ha, 2008, Phys. Rev. Lett. **101**(21), 218001 (pages 4).
Jones, D., 2012, Good practice in (pseudo) random number generation for bioinformatics applications, as of 7 May 2010, available from `http://www0.cs.ucl.ac.uk/staff/d.jones/GoodPracticeRNG.pdf`, accessed 21 Aug 2012.
Juanico, D. E., C. Monterola, and C. Saloma, 2007, Phys. Rev. E **75**(4), 045105(R) (pages 4).
Kadanoff, L. P., S. R. Nagel, L. Wu, and S.-m. Zhou, 1989, Phys. Rev. A **39**(12), 6524.

Kaneko, K., 1989, Physica D **37**(1–3), 60.
Karmakar, R., S. S. Manna, and A. L. Stella, 2005, Phys. Rev. Lett. **94**(8), 088002 (pages 4).
Kernighan, B. W., and R. Pike, 2002, *The Practice of Programming* (Addison-Wesley, Boston, MA, USA).
Kernighan, B. W., and D. M. Ritchie, 1988, *The C programming language* (Prentice Hall, Englewood Cliffs, NJ), 2nd edition.
Klein, W., and J. Rundle, 1993, Phys. Rev. Lett. **71**(8), 1288, comment on (Olami *et al.* 1992), reply (Christensen 1993).
Kloster, M., S. Maslov, and C. Tang, 2001, Phys. Rev. E **63**(2), 026111 (pages 4).
Knuth, D. E., 1997, *The Art of Computer Programming Volumes 1–3* (Addison-Wesley, Reading, MA, USA).
Koplik, J., and H. Levine, 1985, Phys. Rev. B **32**(1), 280.
Landau, D. P., and K. Binder, 2005, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, UK), 2nd edition.
Lauritsen, K. B., S. Zapperi, and H. E. Stanley, 1996, Phys. Rev. E **54**(3), 2483.
Le Doussal, P., K. J. Wiese, and P. Chauve, 2002, Phys. Rev. B **66**(17), 174201 (pages 34).
Lebowitz, J. L., C. Maes, and E. R. Speer, 1990, J. Stat. Phys. **59**, 117, ISSN 0022-4715.
Leschhorn, H., T. Nattermann, S. Stepanow, and L.-H. Tang, 1997, Ann. Physik **6**, 1.
Liggett, T. M., 2005, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes* (Springer-Verlag, Berlin, Germany).
Lin, C.-Y., and C.-K. Hu, 2002, Phys. Rev. E **66**(2), 021307 (pages 12).
Lindman, M., K. Jonsdottir, R. Roberts, B. Lund, and R. Bödvarsso, 2005, Phys. Rev. Lett. **94**(10), 108501 (pages 4), comment (Corral and Christensen 2006).
Lindman, M., K. Jónsdóttir, R. Roberts, B. Lund, and R. Bödvarsso, 2006, Phys. Rev. Lett. **96**(10), 109802 (pages 1), reply to comment (Corral and Christensen 2006).
Lübeck, S., 2000, Phys. Rev. E **61**(1), 204.
Lübeck, S., and P. C. Heger, 2003a, Phys. Rev. E **68**(5), 056102 (pages 11).
Lübeck, S., and P. C. Heger, 2003b, Phys. Rev. Lett. **90**(23), 230601 (pages 4).
Lübeck, S., and K. D. Usadel, 1997, Phys. Rev. E **55**(4), 4095.
Mahieu, S., and P. Ruelle, 2001, Phys. Rev. E **64**(6), 066130 (pages 19).
Majumdar, S. N., and D. Dhar, 1992, Physica A **185**(1–4), 129.
Manna, S. S., 1990, J. Stat. Phys. **59**(1/2), 509.
Manna, S. S., 1991, J. Phys. A: Math. Gen. **24**(7), L363.
Marsaglia, G., 1999, Random numbers for c: The end?, newsgroup posting 20 Jan 1999, 9:00am, accessed 10 May 2011, URL `https://groups.google.com/group/sci.math.num-analysis/msg/eb4ddde782b17051`.
Marsaglia, G., 2005, Mersenne twister, newsgroup posting 14 Jul 2005, 12:04am, available from `http://groups.google.com/group/sci.crypt/browse\_thread/thread/305c507efbe85be4`, accessed 10 May 2011.
Matsumoto, M., 2008, Mersenne twister home page, available from `http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html`, accessed 9 Oct 2008.
Matsumoto, M., and T. Nishimura, 1998, ACM Trans. Model. Comp. Sim. **8**(1), 3.
Middleton, A. A., and C. Tang, 1995, Phys. Rev. Lett. **74**(5), 742.
Milshtein, E., O. Biham, and S. Solomon, 1998, Phys. Rev. E **58**(1), 303.
Mousseau, N., 1996, Phys. Rev. Lett. **77**(5), 968.
Muñoz, M. A., R. Dickman, A. Vespignani, and S. Zapperi, 1999, Phys. Rev. E **59**(5), 6175.
Nakanishi, H., and K. Sneppen, 1997, Phys. Rev. E **55**(4), 4012.

Nattermann, T., S. Stepanow, L.-H. Tang, and H. Leschhorn, 1992, J. Phys. II (France) **2**, 1483.
Newman, M. E. J., and G. T. Barkema, 1999, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, New York, NY, USA).
Olami, Z., H. J. S. Feder, and K. Christensen, 1992, Phys. Rev. Lett. **68**(8), 1244, comment (Klein and Rundle 1993).
Paczuski, M., and K. E. Bassler, 2000a, Phys. Rev. E **62**(4), 5347, according to (Hughes and Paczuski 2002), this(?) version "was published by mistake. The correct version is [...] arXiv:cond-mat/0005340". This probably means arXiv:cond-mat/0005340v2.
Paczuski, M., and K. E. Bassler, 2000b, Theoretical results for sandpile models of soc with multiple topplings, according to (Hughes and Paczuski 2002), this is probably the correct version of (Paczuski and Bassler 2000a)., `arXiv:cond-mat/0005340v2`.
Paczuski, M., and S. Boettcher, 1996, Phys. Rev. Lett. **77**(1), 111, .
Pan, G.-J., D.-M. Zhang, Z.-H. Li, H.-Z. Sun, and Y.-P. Ying, 2005, Phys. Lett. A **338**(3–5), 163 , ISSN 0375-9601.
Pastor-Satorras, R., and A. Vespignani, 2000a, Phys. Rev. E **62**(5), 6195, see (Pastor-Satorras and Vespignani 2000b).
Pastor-Satorras, R., and A. Vespignani, 2000b, J. Phys. A: Math. Gen. **33**(3), L33, see (Pastor-Satorras and Vespignani 2000a).
Patzlaff, H., and S. Trimper, 1994, Phys. Lett. A **189**(3), 187.
Pradhan, P., and D. Dhar, 2006, Phys. Rev. E **73**(2), 021303 (pages 12).
Pradhan, P., and D. Dhar, 2007, J. Phys. A: Math. Theor. **40**(11), 2639.
Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 2007, *Numerical Recipes* (Cambridge University Press, Cambridge, UK), 3rd edition.
Priezzhev, V. B., D. V. Ktitarev, and E. V. Ivashkevich, 1996, Phys. Rev. Lett. **76**(12), 2093.
Privman, V., P. C. Hohenberg, and A. Aharony, 1991, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and J. L. Lebowitz (Academic Press, New York, NY, USA), volume 14, chapter 1, pp. 1–134.
Pruessner, G., 2003, Phys. Rev. E **67**(3), 030301(R) (pages 4).
Pruessner, G., 2012a, The average avalanche size in the manna model and other models of self-organised criticality, arXiv:1208.2069, `arXiv:1208.2069`.
Pruessner, G., 2012b, The field theory of self-organised criticality i: General structure and tree level, in preparation.
Pruessner, G., 2012c, *Self-Organised Criticality* (Cambridge University Press, Cambridge, UK).
Pruessner, G., and H. J. Jensen, 2002a, Phys. Rev. E **65**(5), 056707 (pages 8).
Pruessner, G., and H. J. Jensen, 2002b, Europhys. Lett. **58**(2), 250.
Pruessner, G., and H. J. Jensen, 2004, Phys. Rev. E **70**(6), 066707 (pages 25).
Pruessner, G., and O. Peters, 2006, Phys. Rev. E **73**(2), 025106(R) (pages 4), comment (Alava *et al.* 2008).
Pruessner, G., and O. Peters, 2008, Phys. Rev. E **77**(4), 048102 (pages 2), reply to comment (Alava *et al.* 2008).
Ramasco, J. J., M. A. Muñoz, and C. A. da Silva Santos, 2004, Phys. Rev. E **69**(4), 045105(R) (pages 4).
Rose, G., 2011, Kiss: A bit too simple, preprint from `http://eprint.iacr.org/2011/007.pdf` accessed 20 Aug 2012.
Rossi, M., R. Pastor-Satorras, and A. Vespignani, 2000, Phys. Rev. Lett. **85**(9), 1803.
Ruelle, P., 2002, Phys. Lett. B **539**(1), 172.
Ruelle, P., and S. Sen, 1992, J. Phys. A: Math. Gen. **25**(22), L1257.
Salas, J., and A. D. Sokal, 2000, J. Stat. Phys. **98**(3–4), 551.
Socolar, J. E. S., G. Grinstein, and C. Jayaprakash, 1993, Phys. Rev. E **47**(4), 2366.
Sornette, A., and D. Sornette, 1989, Europhys. Lett. **9**(3), 197.

Sornette, D., and M. J. Werner, 2009, in *Encyclopedia of Complexity and Systems Science*, edited by R. A. Meyers (Springer-Verlag, New York, NY, USA), volume 9, pp. 7872–7891, `arXiv:0803.3756v2`.

Stanley, H. E., 1971, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, New York, NY, USA).

Stauffer, D., and A. Aharony, 1994, *Introduction to Percolation Theory* (Taylor & Francis, London, UK).

Tadić, B., and D. Dhar, 1997, Phys. Rev. Lett. **79**(8), 1519.

Tang, C., and P. Bak, 1988, Phys. Rev. Lett. **60**(23), 2347.

Tebaldi, C., M. De Menech, and A. L. Stella, 1999, Phys. Rev. Lett. **83**(19), 3952.

Torvund, F., and J. Frøyland, 1995, Phys. Scripta **52**, 624.

Turcotte, D. L., 1993, *Fractals and chaos in geology and geophysics* (Cambridge University Press, Cambridge, UK).

Turcotte, D. L., 1999, Rep. Prog. Phys. **62**, 1377.

van der Linden, P., 1994, *Expert C Programming* (Sunsoft Press, A Prentice Hall Title, Mountain View, CA, USA).

van Wijland, F., 2002, Phys. Rev. Lett. **89**(19), 190602 (pages 4).

Černák, J., 2006, Phys. Rev. E **73**, 066125.

Vespignani, A., R. Dickman, M. A. Muñoz, and S. Zapperi, 1998, Phys. Rev. Lett. **81**(25), 5676.

Vespignani, A., R. Dickman, M. A. Muñoz, and S. Zapperi, 2000, Phys. Rev. E **62**(4), 4564.

Vespignani, A., and S. Zapperi, 1995, Phys. Rev. E **51**(3), 1711, see (Lin and Hu 2002).

Wegner, F. J., 1972, Phys. Rev. B **5**(11), 4529.

Welinder, P., G. Pruessner, and K. Christensen, 2007, New J. Phys. **9**(5), 149 (pages 18).

Werner, M. J., and D. Sornette, 2007, Phys. Rev. Lett. **99**(17), 179801, comment on (Davidsen and Paczuski 2005), reply (Davidsen and Paczuski 2007).

Zapperi, S., K. B. Lauritsen, and H. E. Stanley, 1995, Phys. Rev. Lett. **75**(22), 4071.

Zhang, Y.-C., 1989, Phys. Rev. Lett. **63**(5), 470.

Chapter 8
SOC Laboratory Experiments

Gunnar Pruessner

Abstract The experimental verification of Self-Organised Criticality (SOC), or, as in the particular case, its falsification, stands at the very beginning of its long history. Many laboratory experiments have been performed, in diverse areas such as granular media, disordered systems, mechanical instabilities and living systems, some of which have immediate links to *observed* natural phenomena where SOC is widely studied as a model and explanatory framework, such as earthquakes. In the following, after a brief discussion of experimental data analysis methods, a wide range of experiments are presented. Many of them remain ambiguous in the outcome, but a number provide striking evidence for SOC in nature.

8.1 Introduction

The ultimate authority of physical theories is the experiment. Soon after Self-Organised Criticality (SOC) was proposed as an explanation of fractality in nature, laboratory experiments were conducted to probe for its existence in controlled environments.

There is a fine line between conducting an SOC experiment and analysing observational data for the signature of SOC. Given that individual chapters in this book are dedicated for the search for SOC in earthquakes, solar physics, financial markets *etc.*, the following sections focus on experimental setups that can be manipulated and controlled by the experimenter.

One may argue that some experiments addressed questions that were not addressed and not intended to be addressed by the original work by Bak *et al.* (1987). One might even go as far as claiming that the Sandpile Model was

Imperial College London, Department of Mathematics, e-mail: `g.pruessner@imperial.ac.uk`

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

not intended to be a model of sandpiles at all. Rather, it received that name because some of its features resemble some of the physics one might imagine to be relevant in the dynamics of real sandpiles. This point remains open to interpretation and is better left to historians. It is clear, however, that a large number of researchers expected the Sandpile Model, as well as others, such as the Forest Fire Model (Bak *et al.* 1990, Henley 1989, Drossel and Schwabl 1992a) and the Olami-Feder-Christensen (Olami *et al.* 1992), to have some immediate bearing on experiments.

The remainder of this section contains some general remarks of how to identify SOC in experiments and observations (and how not to). This is then followed by four sections about different areas of experimental physics and biology, which have been proved to be a fruitful hunting ground for SOC in nature. Much of this material is based on my earlier review (Pruessner 2012b), which in turn has drawn on a number of reviews, articles and theses by other authors (Turcotte 1993, Carlson *et al.* 1994, Feder 1995, Malcai *et al.* 1997, Turcotte 1999, Gisiger 2001, Dickman *et al.* 2000, Altshuler and Johansen 2004, Kakalios 2005, Wijngaarden *et al.* 2006, Bonachela 2008).

8.1.1 Identifying SOC in experimental data

The numerical and experimental data analysis in the early days of SOC is riddled with technical problems and misunderstandings. They are often caused by assumptions being made at the stage of the data analysis, for example when assuming that the normalisation of a probability density function (PDF) should give rise to a scaling relation (even though it does not, Frette *et al.* 1996, Christensen *et al.* 2008) or by analysing rescaled quantities (even when that spoils the scaling, Planet *et al.* 2009, 2010, Pruessner 2010). This may not come as a surprise given the effort experimenters have to make in order to obtain the raw data. The methods to extract power laws from data remain the subject of an ongoing debate (*e.g.*, Clauset *et al.* 2009).

One of the most damaging problems within the SOC community is the lack of an agreed definition of SOC. In the original spirit of Bak *et al.* (1987) many *theoreticians* would probably agree that SOC systems

- relax **intermittently**, *i.e.* in **avalanches**. More specifically, in SOC systems the time scales of driving and relaxation are (normally not perfectly) separated, with the external drive being very slow compared to the internal relaxation.
- display **non-trivial scale invariance** (as often indicated by power-law distributions and correlations). Non-trivial scaling is normally due to the presence of **non-linearities**, which in SOC are usually caused by the dynamics being subject to thresholds.
- do not require tuning of some **control parameter**. Instead, the system seems to tune itself to the critical point of a continuous phase transition.

As a result, the only scaling parameter is the (necessarily finite) system size, *i.e.* SOC systems display **finite size scaling**.

Ideally, SOC systems have associated with them a continuous phase transition, as originally envisaged. SOC systems are thus truly systems that organise themselves to a critical point without the need of external manipulation, other than the separation of time scales which does not, however, provide an independent scale like a critical temperature in traditional critical phenomena (Stanley 1971).

Much of the confusion in the literature about SOC has been caused by the usage of the term **criticality**. In their original work Bak *et al.* (1987) use it to mean "scale invariance" as found in ordinary phase transitions and critical phenomena. However, and very unfortunately, they also use it to denote the local **thresholds** in the dynamics. That threshold being a maximum slope, the notion of criticality as the system being globally close to the threshold, in a state of maximum susceptibility, marginally stable (Diamond and Hahm 1995), standing at "edge of chaos" (Langton 1990), has become ingrained in the literature. In fact, some authors (*e.g.*, Phillips 2009) consider SOC as a mechanism that optimises the response of the system by pushing to a "tipping point".

However, in general, the local degree of freedom in an SOC is *not* expected to be particularly close to the threshold. SOC systems do not stand at the edge of a precipice in that sense. However, they are (meant to be) critical in the sense of a susceptibility (and thus response) divergent in the system size. For example, in the Manna Model, which bears every feature of a critical, scale-invariant system that self-organises to the critical point, a macroscopic fraction of sites is fully *relaxed*. The system is not critical in the sense of all sites being charged to just below the critical level. If the system is in that stage, the response to an external perturbation is uncharacteristically big and in fact not sustainable by repeated driving and relaxation. In fact, neither correlation nor distribution functions would display broad, algebraic behaviour and moments of event sizes would be trivial functions of the system size.

An SOC system generally does not **optimise** or **maximise** anything beyond the quantities normally associated with singular behaviour at continuous phase transitions. In particular, SOC systems are generally not particularly "efficient" or "tolerant". SOC should not be regarded as the framework to satisfy the demand for such features, as often argued in the context of evolution ("it's SOC because it must evolve to an optimal state").

As for the critical state and the continuous phase transition, as mentioned above, it is often very difficult to identify it in the underlying process. Only very few authors have managed to do so (Peters *et al.* 2002, Peters and Neelin 2006, Neelin *et al.* 2009).

Another feature that is commonly searched for in suspected SOC systems is **1/f noise**. While interesting in its own right, and famously studied for very many years (van der Ziel 1950), it provided the initial motivation for

SOC, but is no longer thought to be at the centre of interest. This is, firstly, because SOC is concerned with long range spatio-temporal correlations which are often much better captured in observables other than the power spectrum. Secondly, a power law power spectrum is not necessarily indicative of power law temporal correlations. Also, such time correlations are often regarded as trivial — in non-equilibrium they are often found generically and thus for rather simple reasons (Grinstein 1995). Finally, intermittent, avalanche-like behaviour, clearly a distinctive feature of SOC systems, is not well captured by $1/f$ noise.

It seems popular to assume that the presence of scaling (in the PDF, a correlation function *etc.*) alone means that SOC is at work. Obviously this is not the case, as demonstrated by the broad variety of traditional critical phenomena. In fact, the initial claim was a different one, namely that SOC might underpin virtually all *spontaneous* occurrence of scaling in nature and, in turn, that slowly driven, intermittent, interaction dominated systems (Jensen 1998) display scaling. This programme, however, has clearly and undoubtedly failed, in computer experiments just as much as in laboratory experiments. The observation of scaling in connection with a *claim* of its cause being SOC does not amount to evidence in its favour.

Finally, it is worth stressing that self-organisation does not necessarily amount to self-organised criticality. In the more recent literature those two concept are sometimes regarded as the same.

8.1.2 Tools and features

To positively determine SOC, as characterised in the list above (Sec. 8.1.1, p. 312), the system considered should display avalanching, non-trivial scale invariance and no external tuning of a control parameter. The avalanching, or, more generally, intermittent behaviour, should be due to a separation of time scales between internal relaxation and external drive, which slowly supplies the system with energy, particles or any other quantity which may be conserved in the bulk but dissipated at the boundaries. This is a qualitative features and fairly easy to determine, with the proviso that any time scale separation is bound to be imperfect in any real-life system.

As for the absence of tuning, this is normally quite obvious and becomes a problem in experiments only when a lot of tweaking, data-detrending and general "clean-up" is necessary to extract the "desired" feature. Many experimentalists rescale histograms so that Gaussian behaviour is easily identified, namely by shifting the observable by its mean and rescaling by its standard deviation. This, however, might obstruct the search for scaling (Pruessner 2010).

The **non-trivial scaling** is the most prominent feature in the data analysis of an SOC system: Non-trivial is to be understood in the non-technical

sense of significant and not expected for obvious reasons, as well as in the technical sense where trivial scaling is found in the absence of **interaction**, for example in Gaussian theories (Cardy 1997). It should be clear from the data and the physics of the system under consideration that interaction between constituent parts are important. It might even be obvious how thresholds come into play.

As for the scaling to be identified in SOC systems, that refers to an **algebraic**, *i.e.***power law relation** between observables. These power laws are parameterised by **exponents**, which are expected to be **universal**. They are therefore the prime candidates to be compared between different systems and models, although many more universal quantities exist.

One reason why SOC is so fascinating is the **emergence** of **effective long-range interaction** that arises from the presence of algebraic correlations — the correlation functions in time and space are not governed by an internal scale, as would be found in the form of a characteristic scale as it enters the archetypal exponential. Rather, they are power laws of distance in time and space, with the dimensional consistency being restored by **metric factors** which are related to microscopic parameters like the resolution of the measuring device. While correlation functions provide the immediate test for emergent behaviour, they are often difficult to obtain and too noisy to be analysed. The typical observable in an SOC study are therefore histograms, which are estimates of the underlying probability density function (PDF). The observables are usually some geometric features (in the widest sense) of the avalanches, such as the avalanche size or extent, but also its duration.

Most experimental systems and numerical models are necessarily finite and therefore the algebraic, scaling regime can only ever be approximate and *intermediate* (Barenblatt 1996), *i.e.* PDFs and correlation functions are cut off sharply at some characteristic scale. *Scaling is an asymptotic phenomenon.* **Finite size scaling** is thus, strictly, the only form of scaling in SOC. This is because the system size is, supposedly, the only constraining feature in SOC that keeps it away from the critical point (also Pruessner 2012a). Although some experiments have been set up so that finite size scaling can be implemented (Held *et al.* 1990, Frette *et al.* 1996), this is often not possible, in particular not when analysing observational data. In these cases, one can introduce cutoffs that implement effectively a finite size through another constraint (*e.g.*, Peters and Neelin 2006) — while it is difficult to make the system effectively bigger, it might be straight-forward to make it effectively smaller. Another, very promising route is **block scaling** (*e.g.*, Pruessner 2008).

Some authors are very strict about the number of decades an observables needs to cover in order to allow for the extraction of a power law (*e.g.*, Malcai *et al.* 1997, Avnir *et al.* 1998). Some expect at least two, better three decades,[1] which is, however, somewhat arbitrary as it depends strongly on

[1] It is popular in the experimental literature to plot the logarithm of the data on linear scales, rather than the original data on a logarithmic scale. This practice makes

the control parameter. Plotting, for example, the PDF of the volume of the avalanches may cover many more decades than plotting the *same* data for the linear extent of avalanches. What matters is the statistics of the data and what can be reliably inferred from it. If only a handful of data points are present per decade (typically 5 to 10), then fitting a power law over little more than one might be futile.

While some very precise methods for fitting power laws have been proposed (Clauset *et al.* 2009), assuming

$$\mathcal{P}(s) = as^{-\tau} \tag{8.1}$$

for the PDF of the avalanche sizes and thus extracting the **(avalanche size) exponent** τ is often inferior to other methods. If finite size scaling applies, or block scaling for that matter, the PDF follows

$$\mathcal{P}(s) = as^{-\tau}\mathcal{G}(s/s_c) \text{ for } s \gg s_0 \ , \tag{8.2}$$

with **metric factor** a and **scaling function** $\mathcal{G}$. The **upper cutoff** s_c is a function (normally to leading order a power law itself) of the system size in case of finite size scaling and of the constraint in case of block scaling. Eq. (8.2) applies for s *sufficiently* large compared to the **lower cutoff** s_0. Because of the presence of the scaling function, Eq. (8.2) is difficult to fit against (but see the discussion on data analysis in Chapter 7), in particular if data is sparse. It might be much easier and much more reliable to perform a moment analysis (De Menech *et al.* 1998), which uses the fact that Eq. (8.2) implies that the nth moment of s, is a power law in s_c,

$$\langle s^n \rangle \simeq as_c^{1+n-\tau} g_n \tag{8.3}$$

with (leading order[2]) amplitude g_n. *Because g_n in Eq. (8.3) and $\mathcal{G}$ in Eq. (8.2) are a priori unknown, one cannot extract the exponent τ without varying s_c.*

The upper cutoff s_c is normally itself a power-law (with corrections) in the system size or any other constraint L,

$$s_c \simeq bL^D \ , \tag{8.4}$$

with metric factor b and avalanche dimension D. Studying therefore the moments as a function of L will therefore produce another universal exponent.

It is striking that in the experimental literature finite size scaling is rarely performed, even when it is at the very heart of SOC. Also, instead of studying moments, most authors are content with the histogram, fitting it in some essentially arbitrary region against a power law and thus extracting an esti-

---

it difficult to extract data for particular values and makes the scale ambiguous when the basis of the logarithm is not stated explicitly.

[2] As with all asymptotic behaviour, sub-leading terms are expected to enter, but are omitted in the present discussion.

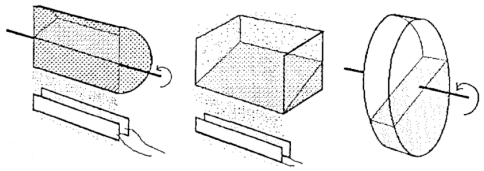


Fig. 8.1: The three different experimental setups used by Jaeger *et al.* (1989). The sand falling out of the open countainers was detetected by the capacitor underneath. Reprinted figure with permission from H. M. Jaeger, C.-h. Liu, and S. R. Nagel, Physical Review Letters, **62**, 40, 1989. Copyright (1989) by the American Physical Society. `http://link.aps.org/abstract/PRL/v62/p40`

mate of the exponent τ. This is all the more regrettable as histograms often come without estimates of their statistical errors. For large, rare events, they can be very large which can, to some extent be controlled by **binning** of the data (see Chapter 7).

8.2 Granular Media

Experiments on $1/f$ noise in granular media, as if in anticipation of SOC decades later, go back to Schick and Verveen (1974). The first experiments to be published in direct response to the original BTW Sandpile Paper was the work by Jaeger *et al.* (1989), also known as the **Chicago group**. As shown in Figure 8.1, their experiments were conducted with three different setups. In each, a **drum** filled with a granular material (glass beads or aluminium-oxide particles) was slowly tilted. As the tilt angle grows beyond the angle of repose, avalanches are triggered and subsequently the granular material may fall out of the drum into a capacitor which determines the volume of the spill by changes in the capacitance. Similar experiments by the group were based on "top loading", whereby grains were slowly added at random to the open surface.

The time series obtained in these early experiments did *not* display the expected scale invariance. Rather, it displayed **oscillatory behaviour** as is sometimes also observed in earthquake data (Schwartz and Coppersmith 1984).[3] Jensen *et al.* (1989) argued that the experiment differed in a number of aspects from the original model: Sand was added continuously in time and space, *i.e.* time scales of driving and relaxation were not perfectly separated. Also, as opposed to the original model, the avalanche size was not measured

[3] The author would like to acknowledge S Hergarten for providing this reference.

as the number of sites toppling (also known as the **flip number**, Kadanoff *et al.* 1989), but as the number of particles falling off the rim, the **drop number**.

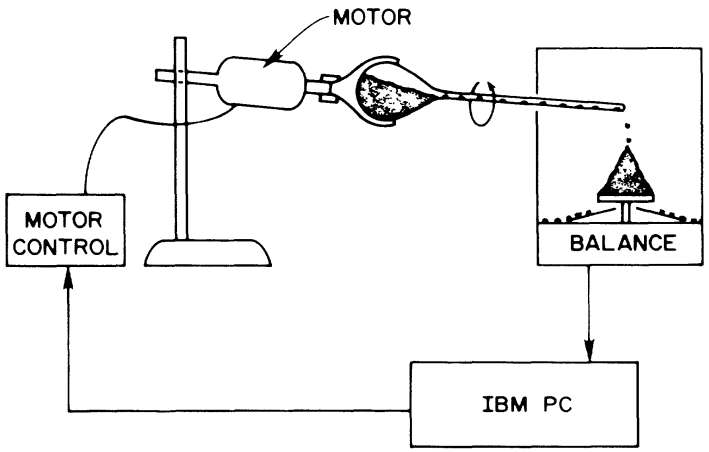


Fig. 8.2: The slowly driven sandpile of Held *et al.* (1990) on the plate of a balance. Avalanches that lead to sand spilling over the edge were detected as a (sudden) decrease in weight, which stopped the feeding mechanism for a suitable amount of time. Figure reprinted with permission from Cambridge University Press and the author (Pruessner, *Self-Organised Criticality*, 2012, Cambridge University Press; Fig 3.2, p. 55).

Held *et al.* (1990), also known as the **IBM group** responded to some of this criticism by implementing a stricter separation of timescales in their experiments on **sandpiles on a balance** and by considering, for the first time, systematic finite size scaling. As can be seen in their experimental setup, Figure 8.2, avalanche sizes were again measured by their drop number, namely by recording the weight of the sandpile as a function of time. A feedback loop allowed the slow sprinkling of the grains to stop whenever an avalanche was running.

The results supported the SOC paradigm to large extent, as the data showed much clearer scaling of the power spectrum of the signal, as well as finite size scaling. However, in sufficiently large systems, the oscillatory behaviour reported earlier was also observed. One may therefore argue that the physics on the large scale is in fact different from that operating on smaller scales (Rosendahl *et al.* 1993).

While the experimental results were encouraging, they also suggested that the expected ubiquity and thus the ease with which SOC should be found in natural and experimental system was an illusion. The slight publication bias

towards what may be perceived as successful experiments (but see for example Jaeger *et al.* 1989, Zieve *et al.* 1996, Nowak *et al.* 1997, Kirchner and Weil 1998) blurs the picture further. Experimentalists had to go to great length to find the desired scaling bahviour, controling, for example, the environmental humidity to avoid unintended cohesion (Nagel 1992, Jaeger and Nagel 1992, Albert *et al.* 1997). Around the same time, however, Evesque (1991) raised inertia as a problem (also Grinstein 1995, Jensen 1998, Dickman *et al.* 2000). In principle, inertia "works both ways", namely by preventing movement as well as sustaining it. However, the thresholds believed to dominate avalanching in sandpiles, such as static friction and steric constraints, are not directly linked to a minimal velocity. Inertia is therefore expected to be much more important when the momentum of oncoming grains helps them to overcome a force threshold. In the light of this competition of inertia and thresholds, one may thus expect cohesion to be a desirable feature.

Bretz *et al.* (1992) introduced two important experimental improvements: Firstly they fixed a layer of sand to the bottom of the tilting box to suppress the unintended instability caused by the reduced friction between sand and container material, a phenomenon studied further for example by Altshuler *et al.* (2001) and Costello *et al.* (2003). Secondly Bretz *et al.* (1992) measured the flip number of the avalanches by means of a CCD camera (used slightly differently by Jánosi and Horváth 1989). This method was later also employed for example by Frette *et al.* (1996) and Aegerter *et al.* (2003).

Grumbacher *et al.* (1993) were still using the drop number but improved Held *et al.*'s (1990) experimental setup by changing the external driving to guarantee that only one bead is delivered by the feeding mechanism at a time. They reported a high degree of universality across different materials, system sizes and covering of the pile's base plate (also Vandewalle *et al.* 1999).

The famous **ricepile experiment** by Frette *et al.* (1996), Figure 8.3, was concerned with granular material with shape anisotropy, as rice grains were dropped, one by one, in the narrow slit between two sheets of Perspex®. In this experiment, again a CCD camera was used to determine the avalanche size in terms of potential energy release. Its resolution also determined the maximum size of the experiment. Although the feeding was slow, it was turned on permanently. The scaling was found to vary with the type of rice used, which was attributed to a difference in surface structure. The ricepile experiment triggered the development of the rice pile model (Christensen *et al.* 1996), now more widely known as the Oslo Model. The original experiment was extended to three dimensions by Aegerter *et al.* (also Aegerter *et al.* 2004a 2004a), producing a data collapse with exponents, however, that do not match those from theoretical models.

Observational data of avalanches in granular media have been analysed for a signature of SOC as well. Notably, the early study by Noever (1993) on snow avalanches suggests the absence of SOC, even though the data is remarkably smooth and broadly distributed (across twelve orders of mag-

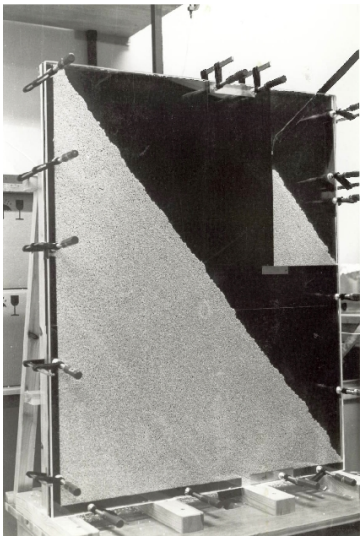


Fig. 8.3: The setup of the ricepile experiment by Frette *et al.* (1996). Two piles (a very large one on the left and a smaller one within the same arrangement towards the right) appear in the narrow slit created by two sheets of Perspex®. Photograph courtesy of V. Frette, K. Christensen, A. Malthe-Sørenssen, J. Feder, T. Jøssang and P. Meakin.

nitude). A later study by Faillettaz *et al.* (2004) supports scaling over two orders of magnitude.

There are a number of reviews of scaling in granular media. Early experimental work was discussed in the review by Feder (1995), concluding that there is little support for SOC. The well-known review by Turcotte (1999) contains a section on granular media, in particular on landslides, as do other more broadly themed reviews, such as the one by Dickman *et al.* (2000) or the relevant chapter by Grinstein (1995) and Bonachela (2008). The reviews by Jaeger *et al.* (1996) and more recently Kakalios (2005) address the physics of granular media in broader terms.

8.3 Systems with internal disorder

Avalanching behaviour in **type II superconductors** was compared very early to sand piles (de Gennes 1966). In these systems, flux lines penetrating the material from the outside are pinned at defects. Because of the repulsion between flux vortices, a defect has an effective carrying capacity. At the **Bean critical state** (Bean 1962), almost all defects are nearly filled com-

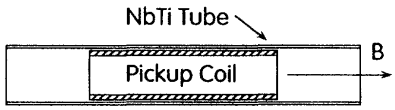


Fig. 8.4: The superconducting SOC experiment by Field *et al.* (1995). A coil in the centre of a (thin) tube of superconducting material registered the change of flux as flux lines arrived there after passing through the superconductor. Reprinted figure with permission from S. Field, J. Witt, F. Nori, and X. Ling, Physical Review Letters, **74**, 1206, 1995. Copyright (1995) by the American Physical Society. `http://link.aps.org/abstract/PRL/v74/p1206`

pletely with flux lines. Further flux lines entering might therefore trigger an avalanche. This phenomenon is sometimes studied from the point of view of interface propagation in random media (Frette 1993, Fisher 1998, Sethna *et al.* 2001). It is closely related to Barkhausen noise discussed below.

The experiments on superconductor avalanches normally use a pickup coil wound around a superconducting specimen which is then exposed to an increasing or decreasing external field, which, for energetic purposes, either forces flux lines to enter or to leave the specimen. The change in flux is picked up by the coil, Figure 8.4.

Pla and Nori (1991) were the first to analyse computer models of superconductivity from an SOC perspective, while Ling *et al.* (1991) were the first to apply it to experimentally obtained current-voltage data, identifying an avalanche-like dynamics and thus confirming the validity of the approach. Wang and Shi (1993) clarified that depending on the strength of th external drive, flux line motion is initially thermally activated before entering an avalanching regime. The experimental setup was further improved by Field *et al.* (1995) who placed the pickup coil at the centre of a tube made of the superconducting material, Figure 8.4. Measuring the flux lines arriving there amounts to measuring the drop number. There are various experimental problems, in particular when flux lines move through the material without significant interaction with the defects or when they intersect the pickup coil only partially.

Zieve *et al.* (1996) were the first to use a Hall probe inside the superconductor, finding no indication of scale invariance in the event statistics. Pla *et al.* (1996) qualified this finding by pointing out that depending on the size and the distribution of pinning centres some materials seem to be more suitable to display SOC-like behaviour than others (also **??**).[4] The negative finding of Zieve *et al.* (1996) was confirmed by Nowak *et al.* (1997) who, in

[4] The author would like to thank Franco Nori for bringing these references to his attention.

addition, found that there was little evidence for avalanching at all in the superconductor. However, the data obtained by Aegerter (1998) using a SQUID (superconducting quantum interference device) displays a broad distribution which seems to be governed by a power law, although only on a narrow range.

Further experimental improvements were made by Behnia *et al.* (2000) using an 8×8 array of Hall probes, which for the first time gave them access to the flip number, *i.e.* the usual measure of the avalanche size in SOC (also Jensen *et al.* 1989). Depending on the temperature range, their findings were compatible with a power-law distribution of events. At low temperatures, avalanches became "catastrophic" moving thousands of flux lines simultaneously.

In 2004 a number of authors published results based on magneto-optical imaging (Vlasko-Vlasov *et al.* 2004, Altshuler *et al.* 2004, Aegerter *et al.* 2004b), all of which confirmed a form of scaling, however with strongly varying exponents. While this might be a matter of material, *i.e.* its physical and chemical composition, and observable, it challenges the assumption of universality, namely that the scaling found in superconductors is merely controlled by its basic microdynamics. However, it is clear that this more recent work on superconductors is one of the best evidence of SOC in experimental systems (reviewed by Altshuler and Johansen 2004, Wijngaarden *et al.* 2006), on a par with granular materials and the observational evidence in rain (Peters *et al.* 2002) and earthquakes (Bak *et al.* 2002, Hergarten 2002).

The **Barkhausen (1919) effect** is another large hunting ground for SOC. It is found in the magnetisation of a ferromagnetic material as it is exposed to a slowly changing external field. Because domains have to re-arrange internally with the changing field, the **pinning and depinning** of their walls causes sudden changes in the magnetisation as do **topological changes** due to merging and sudden appearance and disappearance of domains. The resulting sudden changes in magnetisation are picked up by a coil wound around the sample. Barkhausen noise has a long-standing history, with the distribution of the magnetisation jumps probably first measured by Stierstadt and Boeckh (1965). As early as 1990 Babcock and Westervelt (also Babcock and Westervelt 1989, Babcock *et al.* 1990, Che and Suhl 1990) studied the sudden topological changes in the domain structure using magneto-optical imaging techniques. They found a broad avalanche size distribution which was compatible with a power law. Topological avalanching is generally different from avalanching due to pinning and depinning (Bak and Flyvbjerg 1992, Urbach *et al.* 1995).

Geoffroy and Porteseil (1991a,b, 1994) analysed the power-spectrum of the magnetisation signal, finding $1/f$-noise, strong correlations and further evidence for scaling. Cote and Meisel (1991) confirmed these results, finding even signs for universality. One might therefore be tempted to conclude that Barkhausen noise is a prime example of SOC. This view, however, has been challenged (O'Brien and Weissman 1994, Urbach *et al.* 1995) on the basis of the experimental setup as well as on the basis of the theoretical reasoning

(Alessandro *et al.* 1990, Hardner *et al.* 1993). One widely accept interpretation is that put forward by Perković *et al.* (1995, also Dahmen and Sethna 1993, Sethna *et al.* 1993, Dahmen *et al.* 1994, Dahmen and Sethna 1996, Carrillo *et al.* 1998), namely that a Barkhausen system might simply be *close* to a critical point and therefore display scale invariance over many orders of magnitude without having to self-tune *exactly* to the critical point. These authors studied Barkhausen noise in the context of the random field Ising Model.

Many experiments on **dry friction** and **stick-slip motion** (Heslot *et al.* 1994, for the physics) also belong to the class of systems with internal disorder. The first experiment of this kind was performed by Feder and Feder (1991a) who analysed the stick-slip motion of a piece of sandpaper pulled across a carpet in the light of the **Burridge-Knopoff** Model (Burridge and Knopoff 1967, Olami *et al.* 1992). Earlier experimental work in this context has been (briefly) reviewed by Tullis and Weeks (1986). Feder and Feder (1991a) found the scaling behaviour confirmed over about two orders of magnitude. Ciliberto and Laroche (1994) created a macroscopically rough surface using steel balls embedded in rubber. Again, approximate scaling was found at sufficiently low (driving) speeds. Experiments by **?** on sheared media (with tuning **?**) in a process resembling slow loading of a geological fault, displayed scaling over about one and a half decades. Vallette and Gollub (1993) and Johansen *et al.* (1993) performed friction experiments on materials with disorder on a microscopic scale. While the former were not able to confirm power law behviour, the latter found scaling at sufficiently slow driving speeds, compatible with the separation of time scales necessary in SOC. In more recent experiments, Buldyrev *et al.* (2006) found scaling across almost two orders of magnitude in experiments involving friction of two metal surfaces.

Disorder is also expected to dominate the behaviour often observed of droplets forming and subsequently running across a glass pane (Fisher 1998). Jánosi and Horváth (1989) were the first to analyse such a system in the light of SOC, using an electronic camera and image processing software. They found scale invariance in the power spectrum of the coverage of the pane. Focusing on the ensuing avalanches, Plourde *et al.* (1993) used a similar setup of water droplets forming under in a Perspex® dome. They measured the amount of water arriving at the rim, *i.e.* the drop number, and found scaling governed by exponents which were apparently dependent on the temperature.

8.4 Mechanical instabilities: Fracture and rapture

A number of experiments are concerned with the **acoustic properties** of cracking material. These experiments clearly are to be seen in the context of earthquake models, although cracking is very different from model systems of

interconnected blocks and springs because it is irreversible and the material is destroyed during the course of the experiment.

The literature is somewhat divided as to the extent to which the present phenomena can justifiably be labelled self-organised critical. Two main objections are generally mounted: Firstly, the requirement of *slow* ramping of forces, *i.e.* tuning, and associated with that, the eventual destruction of the material. This argument has also been made in the context of computer models. Given, however, that infinitely slow drive does not provide a finite scale, one might argue that a **separation of time scales** does not constitute a form of tuning of a control parameter. The second objection (terminal destruction) raised, concerns the self-organisation of the material. Provided the material remains intact, cracking and fracturing stops as soon as it supports the imposed stress (or otherwise when it fails completely). Similar to invasion percolation (Wilkinson and Willemsen 1983) this feature might be regarded as a form of extremal dynamics (Miller *et al.* 1993, Cafiero *et al.* 1995, Paczuski *et al.* 1996), which is sometimes itself considered a form of SOC. Further discussions can be found in the theoretical work by Zapperi *et al.* (1997) and Sornette and Andersen (1998). A review with a broader theoretical scope has been published by Sethna (2007), but to large extent those about internal disorder (Frette 1993, Fisher 1998, Sethna *et al.* 2001) also apply in the present context.

One of the first cracking *observations* was undertaken by Diodati *et al.* (1991) who studied the ultrasonic **acoustic emisions** found in volcanic rock below the earth surface in Stromboli (Italy). Stresses are caused by volcanic tremors and explosions and result in ultrasonic sound waves emitted by the surrounding volcanic rock. Typical scaling regions identified in the distribution of event sizes stretch over about one order of magnitude or less.

Cannelli *et al.* (1993) performed a laboratory experiment investigating a somewhat similar phenomenon in hydrogen precipitation in niobium. Microscopic cracks occur as hydrogen solved in the metal changes its chemical composition by forming a hydride as temperatures are slowly lowered. Microscopic cracks occur due to the strain and the resulting stress caused by difference in density of neighbouring domains. Clean power law behaviour of the event size distribution was observed across about two decades.

The ultrasonic sound emitted by cracking plaster under unaxial stress was analysed by Petri *et al.* (1994), who extended their study to the time autocorrelation function of the acoustic signal. Scaling was identified not only in the histograms (with exponents apparently depending on the samples and the experimental conditions) but also in the correlation function, although, unsurprisingly, being somewhat noisier. Remarkably, the authors find that the exponents extracted display universality and compare well for example to those found by Diodati *et al.* (1991) and Cannelli *et al.* (1993).

In their experiment on single crystals of ice, Weiss and Grasso (1997) studied the acoustic emission of a cylindrical ice crystal subject to uniaxial stress as well as torsion (shear). Although statistics for large events is sparse,

the scaling found covers around three orders of magnitude for uniaxial stress. As opposed to fracturing experiments, the structural changes in ice are due to creep rather than cracks (see also the *plastic instabilities* discussed by **?**). The authors extracted the same exponents for different levels of stress.

Garcimartín *et al.* (1997) identified a scaling region over more than two decades in the acoustic energy emission of cracking plaster, wood and fibreglass prior to breakup over a period of about 2 hours. They argued that the need to slowly increase the pressure in the material (rather than the strain) might well be responsible for the occurrence of power laws and thus were reluctant to label the phenomenon self-organised critical (Guarino *et al.* 1998). However, the material straining to a level to support the imposed pressure might be seen precisely as a process of self-organisation.

Maes *et al.* (1998) analysed their fracturing experiments on cellular glass (a type of foamy glass) along these lines, pointing out in particular that a separation of time scales between driving (over about 100 hours) and relaxation takes place. Clean power law behaviour was found, with exponents robust against a change of the time window considered.

A very different experiment, but one that attracted a lot of attention, is the one on paper crumpling by Houle and Sethna (1996). Again, the acoustic signal was analysed, displaying approximate scaling over about three orders of magnitude in the event size distribution, with no evidence for a systematic dependence on experimental parameters (paper size and method of crumpling). Around the same time, similar experiments had been performed on Mylar® by Kramer and Lobkovsky (1996) who found significant wear of the material and scaling of the event size histogram with an exponent around 1, clearly less than the result by Houle and Sethna (1996). Kramer and Lobkovsky (1996) also studied the energy autocorrelation function, which displays some noisy long-time scaling behaviour. One may argue that the experiments on paper should feature in the section on systems with internal disorder, Sec. 8.3, but this is more difficult to uphold in case of the crumpling of Mylar. Experiments on paper *tearing* performed by Salminen *et al.* (2002) also found scaling (with an exponent depending on the strain rate), but seems to have avoided the discussion in relation to SOC altogether.

8.5 Biological systems

The difference between laboratory experiments and analysis of observational data is probably most blurred in the case of biological systems, which are normally too *complex* to be controlled and tempered with, let alone built from scratch, yet might be studied on a laboratory bench. Chialvo and Bak (1999) promoted the idea that some brain function (learning, plasticity) can be explained by features that are typical for systems displaying SOC. A number of researchers have addressed the open questions in experiments.

At least three extensive reviews are available on the subject area (Stanley *et al.* 1996, Brown and West 2000, Gisiger 2001), which focus on scaling and scale invariance in living systems, yet not on SOC. Evidence for power-law correlations "in the brain" has been collated by Chialvo (2004) and more recently by Chialvo *et al.* (2008).

The first vast area to mention in the present context is that of neural networks (Tononi and Edelman 1998). Starting with the theoretical overview Hopfield (1994), the first to study neural networks from the point of view of SOC were Corral *et al.* (1995), Chen *et al.* (1995) and Herz and Hopfield (1995). Although some authors discussed self-similarity and scaling in neural systems (Teich *et al.* 1997, Lowen *et al.* 1997, Toib *et al.* 1998) the first test for SOC in experimental data was done by Papa and da Silva (1997) based on the results by Gattass and Desimone (1996), who studied the activity of **cortex neurons** of macaques. Worrell *et al.* (2002) used data obtained from human brains to make the case for SOC in an epileptic brain and Bédard *et al.* (2006) developed a model of brain platicity on the basis of SOC.

Probably because of the original work by Bak *et al.* (1987), Linkenkaer-Hansen *et al.* (2001) used $1/f$ noise as an indicator for the presence of SOC (as found earlier by Teich *et al.* 1997). This view was later criticised by Bédard *et al.* (2006, also Mazzoni *et al.* 2007).

Segev *et al.* (2002) were the first to study spatiotemporal patterns in neuronal networks, however not with an SOC perspective, even when many of the histograms found display scaling over up to two orders of magnitude. The phrase of "neural avalanches" (possible inspired by the "earthquakes in the brain" as of Papa and da Silva 1997) was coined in the much-celebrated work by Beggs and Plenz (2003), who considered the critical state as one with optimal information transmission and processing (Beggs 2008).

Rather than studying the physiological properties of the brain, some researchers focused on the behavioural, psychological side. Van Orden *et al.* (2003) identified SOC in human cognition, again on the basis of $1/f$ noise. The experimental basis was a reaction-time task (the time series being the sequence of 1100 reaction times to an event) and a word-naming task (series of 1100 pronounciations, *i.e.* namings, of randomly chosen four- or five-letter words appearing on a screen). Although the data is bound to be noisy and sparse, there is indication of scaling. The analysis and its conclusion was heavily criticised by Wagenmakers *et al.* (2005, response: Van Orden *et al.* 2005), who also tested the data against a range of standard statistical models.

In particular in biology, while evidence for scaling exists it does not always seem possible to identify the functional role of avalanches or, more specifically, SOC (Mazzoni *et al.* 2007). At least two studies use SOC at most as a point of view rather than an explanatory framework. In fact, Tsumiyama *et al.* (2009) seem to use the term to mean "maximum sustainable level before breakdown". Similarly, Phillips (2009) refers to it as a maximisation principle in an analysis of protein properties (Feder *et al.* 1984, Jossang *et al.* 1985, Moret and Zebende 2007).

On the other hand, it is clear that much of the data obtained in experiments on living systems does not lend itself naturally to an extensive analysis, probing for critical behaviour and finite size scaling as described in the introduction (p. 315). For example, the extensive data analysis performed by Moret and Zebende (2007) on amino acids does not go much beyond a decade. One may wonder whether a concept like SOC can be at work in a system where it cannot be probed for, not even in principle.

Acknowledgements The author gratefully acknowledges the kind support by EPSRC Mathematics Platform grant EP/I019111/1.

References

Aegerter, C. M., 1998, Phys. Rev. E **58**(2), 1438.
Aegerter, C. M., R. Günther, and R. J. Wijngaarden, 2003, Phys. Rev. E **67**(5), 051306 (pages 6).
Aegerter, C. M., K. A. Lőrincz, M. S. Welling, and R. J. Wijngaarden, 2004a, Phys. Rev. Lett. **92**(5), 058702 (pages 4).
Aegerter, C. M., M. S. Welling, and R. J. Wijngaarden, 2004b, Europhys. Lett. **65**(6), 753.
Albert, R., I. Albert, D. Hornbaker, P. Schiffer, and A.-L. Barabási, 1997, Phys. Rev. E **56**(6), R6271.
Alessandro, B., C. Beatrice, G. Bertotti, and A. Montorsi, 1990, J. Appl. Phys. **68**(6), 2901.
Altshuler, E., and T. H. Johansen, 2004, Rev. Mod. Phys. **76**(2), 471.
Altshuler, E., T. H. Johansen, Y. Paltiel, P. Jin, K. E. Bassler, O. Ramos, Q. Y. Chen, G. F. Reiter, E. Zeldov, and C. W. Chu, 2004, Phys. Rev. B **70**(14), 140505(R) (pages 4), early preprint `arXiv:cond-mat/0208266`.
Altshuler, E., O. Ramos, C. Martínez, L. E. Flores, and C. Noda, 2001, Phys. Rev. Lett. **86**(24), 5490.
Avnir, D., O. Biham, D. Lidar, and O. Malcai, 1998, Science **279**(5347), 39.
Babcock, K. L., R. Seshadri, and R. M. Westervelt, 1990, Phys. Rev. A **41**(4), 1952.
Babcock, K. L., and R. M. Westervelt, 1989, Phys. Rev. A **40**(4), 2022.
Babcock, K. L., and R. M. Westervelt, 1990, Phys. Rev. Lett. **64**(18), 2168.
Bak, P., K. Chen, and C. Tang, 1990, Phys. Lett. A **147**(5–6), 297.
Bak, P., K. Christensen, L. Danon, and T. Scanlon, 2002, Phys. Rev. Lett. **88**(17), 178501 (pages 4).
Bak, P., and H. Flyvbjerg, 1992, Phys. Rev. A **45**(4), 2192.
Bak, P., C. Tang, and K. Wiesenfeld, 1987, Phys. Rev. Lett. **59**(4), 381.
Barenblatt, G. I., 1996, *Scaling, self-similarity, and intermediate asymptotics* (Cambridge University Press, Cambridge, UK).
Barkhausen, H., 1919, Physik. Zeitschr. **20**, 401.
Bean, C. P., 1962, Phys. Rev. Lett. **8**(6), 250.
Bédard, C., H. Kröger, and A. Destexhe, 2006, Phys. Rev. Lett. **97**(11), 118102 (pages 4).
Beggs, J. M., 2008, Phil. Trans. R. Soc. A **366**(1864), 329.
Beggs, J. M., and D. Plenz, 2003, J. Neurosci. **23**(35), 11167.
Behnia, K., C. Capan, D. Mailly, and B. Etienne, 2000, Phys. Rev. B **61**(6), R3815.

Bonachela, J. A., 2008, *Universality in Self-Organized Criticality*, Ph.D. thesis, Departmento de Electromagnetismo y Fisíca de la Materia & Institute Carlos I for Theoretical and Computational Physics, University of Granada, Granada, Spain, accessed 12 Sep 2009, URL `http://hera.ugr.es/tesisugr/17706312.pdf`.

Bretz, M., J. B. Cunningham, P. L. Kurczynski, and F. Nori, 1992, Phys. Rev. Lett. **69**(16), 2431.

Brown, J. H., and G. B. West (eds.), 2000, *Scaling in Biology*, Santa Fe Institute Studies on the Sciences of Complexity (Oxford University Press, New York, NY, USA).

Buldyrev, S. V., J. Ferrante, and F. R. Zypman, 2006, Phys. Rev. E **74**(6), 066110 (pages 12).

Burridge, R., and L. Knopoff, 1967, Bull. Seismol. Soc. Am. **57**(3), 341.

Cafiero, R., V. Loreto, L. Pietronero, A. Vespignani, and S. Zapperi, 1995, Europhys. Lett. **29**(2), 111.

Cannelli, G., R. Cantelli, and F. Cordero, 1993, Phys. Rev. Lett. **70**(25), 3923.

Cardy, J., 1997, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, Cambridge, UK).

Carlson, J. M., J. S. Langer, and B. E. Shaw, 1994, Rev. Mod. Phys. **66**(2), 657.

Carrillo, L., L. Mañosa, J. Ortín, A. Planes, and E. Vives, 1998, Phys. Rev. Lett. **81**(9), 1889.

Che, X., and H. Suhl, 1990, Phys. Rev. Lett. **64**(14), 1670.

Chen, D.-M., S. Wu, A. Guo, and Z. R. Yang, 1995, J. Phys. A: Math. Gen. **28**(18), 5177.

Chialvo, D., and P. Bak, 1999, Neuroscience **90**(4), 1137 .

Chialvo, D. R., 2004, Physica A **340**(4), 756, proceedings of the symposium Complexity and Criticality: in memory of Per Bak (1947–2002), Copenhagen, Denmark, Aug 21–23, 2003.

Chialvo, D. R., P. Balenzuela, and D. Fraiman, 2008, in *Collective Dynamics: Topics on Competition and Cooperation in the Biosciences*, edited by L. M. Ricciardi, A. Buonocore, and E. Pirozzi, volume 1028 of *AIP Conference Proceedings*, pp. 28–45, proceedings of the International Conference BIOCOMP2007 on Collective Dynamics - Topics on Competition and Cooperation in the Biosciences, Vietri, Italy, Sep 24–28, 2007, `arXiv:0804.0032`.

Christensen, K., Á. Corral, V. Frette, J. Feder, and T. Jøssang, 1996, Phys. Rev. Lett. **77**(1), 107.

Christensen, K., N. Farid, G. Pruessner, and M. Stapleton, 2008, Eur. Phys. J. B **62**(3), 331.

Ciliberto, S., and C. Laroche, 1994, J. Phys. I (France) **4**, 223.

Clauset, A., C. R. Shalizi, and M. E. J. Newman, 2009, SIAM Rev. **51**(4), 661.

Corral, Á., C. J. Pérez, A. Díaz-Guilera, and A. Arenas, 1995, Phys. Rev. Lett. **74**(1), 118.

Costello, R. M., K. L. Cruz, C. Egnatuk, D. T. Jacobs, M. C. Krivos, T. S. Louis, R. J. Urban, and H. Wagner, 2003, Phys. Rev. E **67**(4), 041304 (pages 9).

Cote, P. J., and L. V. Meisel, 1991, Phys. Rev. Lett. **67**(10), 1334.

Dahmen, K., S. Kartha, J. A. Krumhansl, B. W. Roberts, J. P. Sethna, and J. D. Shore, 1994, J. Appl. Phys. **75**(10), 5946, proceedings of the 38th Annual Conference On Magnetism and Magnetic Materials, Minneapolis, MN, USA, Nov 15–18, 1993.

Dahmen, K., and J. P. Sethna, 1993, Phys. Rev. Lett. **71**(19), 3222.

Dahmen, K., and J. P. Sethna, 1996, Phys. Rev. B **53**(22), 14872.

de Gennes, P. G., 1966, *Superconductivity of Metals and Alloys* (Physica B, Reading, MA, USA), translated by P. A. Pincus.

De Menech, M., A. L. Stella, and C. Tebaldi, 1998, Phys. Rev. E **58**(3), R2677.

Diamond, P. H., and T. S. Hahm, 1995, Phys. Plasmas **2**(10), 3640.

Dickman, R., M. A. Muñoz, A. Vespignani, and S. Zapperi, 2000, Braz. J. Phys. **30**(1), 27.
Diodati, P., F. Marchesoni, and S. Piazza, 1991, Phys. Rev. Lett. **67**(17), 2239.
Drossel, B., and F. Schwabl, 1992a, Phys. Rev. Lett. **69**(11), 1629, largely identical to proceedings article (Drossel and Schwabl 1992b).
Drossel, B., and F. Schwabl, 1992b, Physica A **191**(1–4), 47, proceedings of the International Conference on Fractals and Disordered Systems, Hamburg, Germany, Jul 29–31, 1992.
Evesque, P., 1991, Phys. Rev. A **43**(6), 2720.
Faillettaz, J., F. Louchet, and J.-R. Grasso, 2004, Phys. Rev. Lett. **93**(20), 208001 (pages 4).
Feder, H. J. S., and J. Feder, 1991a, Phys. Rev. Lett. **66**(20), 2669, see erratum (Feder and Feder 1991b).
Feder, H. J. S., and J. Feder, 1991b, Phys. Rev. Lett. **67**(2), 283, erratum.
Feder, J., 1995, Fractals **3**(3), 431.
Feder, J., T. Jøssang, and E. Rosenqvist, 1984, Phys. Rev. Lett. **53**, 1403.
Field, S., J. Witt, F. Nori, and X. Ling, 1995, Phys. Rev. Lett. **74**(7), 1206.
Fisher, D. S., 1998, Phys. Rep. **301**(1–3), 113.
Frette, V., 1993, Phys. Rev. Lett. **70**(18), 2762.
Frette, V., K. Christensen, A. Malthe-Sørenssen, J. Feder, T. Jøssang, and P. Meakin, 1996, Nature **379**, 49.
Garcimartín, A., A. Guarino, L. Bellon, and S. Ciliberto, 1997, Phys. Rev. Lett. **79**(17), 3202.
Gattass, R., and R. Desimone, 1996, Rev. Bras. Biol. **56**(Su 1 Pt 2), 257.
Geoffroy, O., and J. L. Porteseil, 1991a, J. Magn. Magn. Mater **97**(1–3), 205.
Geoffroy, O., and J. L. Porteseil, 1991b, J. Magn. Magn. Mater **97**(1–3), 198.
Geoffroy, O., and J. L. Porteseil, 1994, J. Magn. Magn. Mater **133**(1–3), 1.
Gisiger, T., 2001, Biol. Rev. **76**, 161.
Grinstein, G., 1995, in *Scale Invariance, Interfaces, and Non-Equilibrium Dynamics*, edited by A. McKane, M. Droz, J. Vannimenus, and D. Wolf (Plenum Press, New York, NY, USA), pp. 261–293, NATO Advanced Study Institute on Scale Invariance, Interfaces, and Non-Equilibrium Dynamics, Cambridge, UK, Jun 20–30, 1994.
Grumbacher, S. K., K. M. McEwen, D. A. Halverson, D. T. Jacobs, and J. Lindner, 1993, Am. J. Phys. **61**(4), 329.
Guarino, A., A. Garcimartín, and S. Ciliberto, 1998, Eur. Phys. J. B **6**(1), 13.
Hardner, H. T., M. B. Weissman, M. B. Salamon, and S. S. P. Parkin, 1993, Phys. Rev. B **48**(21), 16156.
Held, G. A., D. H. Solina, H. Solina, D. T. Keane, W. J. Haag, P. M. Horn, and G. Grinstein, 1990, Phys. Rev. Lett. **65**(9), 1120.
Henley, C. L., 1989, Bull. Am. Phys. Soc. **34**(3), 838, abstract of talk M18.2, 23 March 1989, of the 1989 March Meeting of The American Physical Society, St. Louis, MO, USA, Mar 20–24, 1989.
Hergarten, S., 2002, *Self-Organized Criticality in Earth Systems* (Springer-Verlag, Berlin, Germany).
Herz, A. V. M., and J. J. Hopfield, 1995, Phys. Rev. Lett. **75**(6), 1222.
Heslot, F., T. Baumberger, B. Perrin, B. Caroli, and C. Caroli, 1994, Phys. Rev. E **49**, 4973.
Hopfield, J. J., 1994, Phys. Today **47**(2), 40, special issue: Physics and Biology.
Houle, P. A., and J. P. Sethna, 1996, Phys. Rev. E **54**(1), 278.
Jaeger, H. M., C.-h. Liu, and S. R. Nagel, 1989, Phys. Rev. Lett. **62**(1), 40.
Jaeger, H. M., and S. R. Nagel, 1992, Science **255**(5051), 1523.
Jaeger, H. M., S. R. Nagel, and R. P. Behringer, 1996, Rev. Mod. Phys. **68**(4), 1259.
Jánosi, I. M., and V. K. Horváth, 1989, Phys. Rev. A **40**(9), 5232.

Jensen, H. J., 1998, *Self-Organized Criticality* (Cambridge University Press, New York, NY, USA).
Jensen, H. J., K. Christensen, and H. C. Fogedby, 1989, Phys. Rev. B **40**(10), 7425.
Johansen, A., P. Dimon, C. Ellegaard, J. S. Larsen, and H. H. Rugh, 1993, Phys. Rev. E **48**(6), 4779.
Jossang, T., J. Feder, and E. Rosenqvist, 1985, The Journal of Chemical Physics **82**(1), 574.
Kadanoff, L. P., S. R. Nagel, L. Wu, and S.-m. Zhou, 1989, Phys. Rev. A **39**(12), 6524.
Kakalios, J., 2005, Am. J. Phys. **73**(1), 8.
Kirchner, J. W., and A. Weil, 1998, Nature **395**(6700), 337.
Kramer, E. M., and A. E. Lobkovsky, 1996, Phys. Rev. E **53**, 1465.
Langton, C. G., 1990, Physica D **42**(1–3), 12.
Ling, X. S., D. Shi, and J. L. Budnick, 1991, Physica C **185–189**(Part 4), 2181.
Linkenkaer-Hansen, K., V. V. Nikouline, J. M. Palva, and R. J. Ilmoniemi, 2001, J. Neurosci. **21**(4), 1370.
Lowen, S. B., S. S. Cash, M.-m. Poo, and M. C. Teich, 1997, J. Neurosci. **17**(15), 5666.
Maes, C., A. Van Moffaert, H. Frederix, and H. Strauven, 1998, Phys. Rev. B **57**(9), 4987.
Malcai, O., D. A. Lidar, O. Biham, and D. Avnir, 1997, Phys. Rev. E **56**(3), 2817.
Mazzoni, A., F. D. Broccard, E. Garcia-Perez, P. Bonifazi, M. E. Ruaro, and V. Torre, 2007, PLoS ONE **2**(5), e439.
Miller, S. L., W. M. Miller, and P. J. McWhorter, 1993, J. Appl. Phys. **73**(6), 2617.
Moret, M. A., and G. F. Zebende, 2007, Phys. Rev. E **75**, 011920.
Nagel, S. R., 1992, Rev. Mod. Phys. **64**(1), 321.
Neelin, J. D., O. Peters, and K. Hales, 2009, J. Atmos. Sci. **66**(8), 2367.
Noever, D. A., 1993, Phys. Rev. E **47**(1), 724.
Nowak, E. R., O. W. Taylor, L. Liu, H. M. Jaeger, and T. I. Selinder, 1997, Phys. Rev. B **55**(17), 11702.
O'Brien, K. P., and M. B. Weissman, 1994, Phys. Rev. E **50**(5), 3446.
Olami, Z., H. J. S. Feder, and K. Christensen, 1992, Phys. Rev. Lett. **68**(8), 1244.
Paczuski, M., S. Maslov, and P. Bak, 1996, Phys. Rev. E **53**(1), 414.
Papa, A. R. R., and L. da Silva, 1997, Theor. Biosci. **116**, 321.
Perković, O., K. Dahmen, and J. P. Sethna, 1995, Phys. Rev. Lett. **75**(24), 4528.
Peters, O., C. Hertlein, and K. Christensen, 2002, Phys. Rev. Lett. **88**(1), 018701 (pages 4).
Peters, O., and J. D. Neelin, 2006, Nat. Phys. **2**(6), 393.
Petri, A., G. Paparo, A. Vespignani, A. Alippi, and M. Costantini, 1994, Phys. Rev. Lett. **73**(25), 3423.
Phillips, J. C., 2009, Phys. Rev. E **80**, 051916.
Pla, O., and F. Nori, 1991, Phys. Rev. Lett. **67**(7), 919.
Pla, O., N. K. Wilkin, and H. J. Jensen, 1996, Europhys. Lett. **33**(4), 297.
Planet, R., S. Santucci, and J. Ortín, 2009, Phys. Rev. Lett. **102**(9), 094502 (pages 4), comment (Pruessner 2010).
Planet, R., S. Santucci, and J. Ortín, 2010, Phys. Rev. Lett. **105**(2), 029402 (pages 1), reply to comment (Pruessner 2010).
Plourde, B., F. Nori, and M. Bretz, 1993, Phys. Rev. Lett. **71**(17), 2749.
Pruessner, G., 2008, New J. Phys. **10**(11), 113003 (pages 13).
Pruessner, G., 2010, Phys. Rev. Lett. **105**(2), 029401 (pages 1), comment on (Planet *et al.* 2009), reply (Planet *et al.* 2010).
Pruessner, G., 2012a, The average avalanche size in the manna model and other models of self-organised criticality, `arXiv:1208.2069`.

Pruessner, G., 2012b, *Self-Organised Criticality* (Cambridge University Press, Cambridge, UK).
Rosendahl, J., M. Vekić, and J. Kelley, 1993, Phys. Rev. E **47**(2), 1401.
Salminen, L. I., A. I. Tolvanen, and M. J. Alava, 2002, Phys. Rev. Lett. **89**, 185503.
Schick, K. L., and A. A. Verveen, 1974, Nature **251**(5476), 599.
Schwartz, D. P., and K. J. Coppersmith, 1984, J. Geophys. Res. **89**(B7), 5681, ISSN 0148-0227.
Segev, R., M. Benveniste, E. Hulata, N. Cohen, A. Palevski, E. Kapon, Y. Shapira, and E. Ben-Jacob, 2002, Phys. Rev. Lett. **88**(11), 118102 (pages 4).
Sethna, J. P., 2007, in *Complex Systems*, edited by M. M. Jean-Philippe Bouchaud and J. Dalibard (Elsevier), volume 85 of *Les Houches*, pp. 257–288, `cond-mat/0612418v1`.
Sethna, J. P., K. Dahmen, S. Kartha, J. A. Krumhansl, B. W. Roberts, and J. D. Shore, 1993, Phys. Rev. Lett. **70**(21), 3347.
Sethna, J. P., K. A. Dahmen, and C. R. Myers, 2001, Nature **410**(6825), 242.
Sornette, D., and J. V. Andersen, 1998, Eur. Phys. J. B **1**(3), 353.
Stanley, H. E., 1971, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, New York, NY, USA).
Stanley, H. E., L. A. N. Amaral, S. V. Buldyrev, A. L. Goldberger, S. Havlin, B. T. Hyman, H. Leschhorn, P. Maass, H. A. Makse, C. K. Peng, M. A. Salinger, M. H. R. Stanley, *et al.*, 1996, Fractals **4**(3), 427.
Stierstadt, K., and W. Boeckh, 1965, Z. Phys. **186**(2), 154.
Teich, M. C., C. Heneghan, S. B. Lowen, T. Ozaki, and E. Kaplan, 1997, J. Opt. Soc. Am. A **14**(3), 529.
Toib, A., V. Lyakhov, and S. Marom, 1998, J. Neurosci. **18**(5), 1893.
Tononi, G., and G. M. Edelman, 1998, Science **282**(5395), 1846.
Tsumiyama, K., Y. Miyazaki, and S. Shiozawa, 2009, PLOS ONE **4**(12).
Tullis, T. E., and J. D. Weeks, 1986, Pure Appl. Geophys. **124**(3), 383.
Turcotte, D. L., 1993, *Fractals and chaos in geology and geophysics* (Cambridge University Press, Cambridge, UK).
Turcotte, D. L., 1999, Rep. Prog. Phys. **62**, 1377.
Urbach, J. S., R. C. Madison, and J. T. Markert, 1995, Phys. Rev. Lett. **75**(2), 276.
Vallette, D. P., and J. P. Gollub, 1993, Phys. Rev. E **47**(2), 820.
van der Ziel, A., 1950, Physica **16**(4), 359.
Van Orden, G. C., J. G. Holden, and M. T. Turvey, 2003, J. Exp. Psychol. Gen. **132**(3), 331.
Van Orden, G. C., J. G. Holden, and M. T. Turvey, 2005, J. Exp. Psychol. Gen. **134**(1), 117.
Vandewalle, N., R. D'hulst, and M. Ausloos, 1999, Phys. Rev. E **59**, 631.
Vlasko-Vlasov, V. K., U. Welp, V. Metlushko, and G. W. Crabtree, 2004, Phys. Rev. B **69**(14), 140504(R) (pages 4).
Wagenmakers, E.-J., S. Farrell, and R. Ratcliff, 2005, J. Exp. Psychol. Gen. **134**(1), 108.
Wang, Z., and D. Shi, 1993, Phys. Rev. B **48**(13), 9782.
Weiss, J., and J.-R. Grasso, 1997, J. Phys. Chem. B **101**(32), 6113.
Wijngaarden, R. J., M. S. Welling, C. M. Aegerter, and M. Menghini, 2006, Eur. Phys. J. B **50**(1), 117.
Wilkinson, D., and J. F. Willemsen, 1983, J. Phys. A: Math. Gen. **16**(14), 3365.
Worrell, G. A., S. D. Cranstoun, J. Echauz, and B. Litt, 2002, NeuroReport **13**(16), 2017.
Zapperi, S., A. Vespignani, and H. E. Stanley, 1997, Nature **388**(6643), 658.
Zieve, R. J., T. F. Rosenbaum, H. M. Jaeger, G. T. Seidler, G. W. Crabtree, and U. Welp, 1996, Phys. Rev. B **53**(17), 11849.

Chapter 9
Self-Organizing Complex Earthquakes: Scaling in Data, Models, and Forecasting

Michael K. Sachs, John B. Rundle, James R. Holliday, Joseph Gran, Mark Yoder, Donald L. Turcotte and William Graves

Abstract In many natural and social systems, observations of the frequency-size distributions of events can be approximated by power-law (scaling or fractal) distributions. An important question is whether the probability of extreme events can be estimated by extrapolating the power-law distributions, or whether the largest events occur more frequently than the power-law extrapolation would predict. In this paper we consider specifically earthquake systems, and discuss both the data supporting these ideas, as well as simple models. Extrapolations using power-laws are widely used in probabilistic hazard assessment. As a result, the existence of events not falling on the scaling line will be misrepresented by these simple models. We also discuss earth-

Michael K. Sachs
Department of Physics, University of California, Davis, CA e-mail: mksachs@ucdavis.edu

John B. Rundle
Department of Physics, University of California, Davis, CA e-mail: jbrundle@ucdavis.edu

James R. Holliday
Department of Physics, University of California, Davis, CA e-mail: jrholliday@ucdavis.edu

Joseph Gran
Department of Physics, University of California, Davis, CA e-mail: jdgran@ucdavis.edu

Mark Yoder
Department of Physics, University of California, Davis, CA e-mail: mryoder@ucdavis.edu

Donald L. Turcotte
Department of Geology, University of California, Davis, CA e-mail: dlturcotte@ucdavis.edu

William Graves
Open Hazards Group, Davis, CA e-mail: graveswr@gmail.com

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

quake forecast models that arise from scaling ideas, specifically the Natural Time Weibull (NTW) model. This idea is based on the premise that the scaling distribution is eventually filled in by smaller events following a previous large event. We illustrate these results by direct application to California and Japan.

9.1 Introduction

Extreme events are large fluctuations that occur suddenly against a background of smaller events. Details of the nature of such events have been discussed recently in Sachs et al. (2012). The statistics of extreme events have been studied extensively (Kinnison 1985). Applications include floods, wars, financial crashes, and many others. Extreme events that change global society were characterized as black swans by Taleb (2007). The frequency-size distributions associated with many natural hazards satisfy power-law (fractal) statistics to a good approximation (Turcotte 1997). They have also been cited as examples of Self-Organized Criticality (SOC). Examples include earthquakes, landslides, volcanic eruptions, and wildfires. In the financial markets, the largest events correspond to financial crashes of markets, currency valuations, and bankruptcy of the largest corporations or financial institutions (Mantegna & Stanley 2000, Sornette 2004). In this paper, we are interested not so much in the scaling behavior itself, but in the departures from pure scaling or SOC.

Forecasting future events generally relies on pattern matching to previous histories of events. An important question in probabilistic hazard assessment is whether future extreme events can be forecast by extrapolating the power-law behavior, and whether the largest outliers recur in some kind of predictable fashion.

In some cases small events satisfy power-law scaling but one or more extreme events are significantly larger than the extrapolation of the power-law scaling. Sornette (2009) refers to this class of extreme events as dragon kings, but they may be more commonly known as nucleation events, or first order phase transitions. Examples might be the population of London and Paris relative to the power-law distribution of the population of the other cities in the U.K. and France. Another example is material fracture. When stress on a brittle material is increased there is often a power-law distribution of acoustic emissions before an unstable fracture propagates through the sample (Guarino et al. 1998). The precursory emissions are a nucleation process and the fracture can be classified as a phase change (Alava et al. 2006).

We will address the question: Are these off-scaling events relevant to probabilistic hazard assessment, or are they inherently unpredictable? Our discussion will emphasize seismic hazard and to a lesser extent, financial hazard.

On relatively large geographic scales the frequency-rupture-area statistics of earthquakes are well represented by power-law statistics even for the largest events. However, this behavior does not seem to be the case locally. On major faults, such as the San Andreas in California, seismic activity is dominated by great earthquakes, like the 1906 San Francisco earthquake. These are known as "characteristic" earthquakes. As a specific example we will consider a characteristic earthquake cycle on the Parkfield segment of the San Andreas fault. We will then consider in some detail the use of slider-block models to represent seismicity.

In the financial area, the tail of the distribution of market moves in the Standard and Poors 500 index is represented by a power-law. The distribution itself is known to be leptokurtotic (Mantegna & Stanley 2000), meaning that the central part of the statistical distribution of price moves is higher than the corresponding Gaussian distribution, but the tails are fatter. Some of the largest moves, such as the famous market crash of October 19, 1987, are outliers on the general trend of power-law scaling (Sornette 2004). They are therefore candidates for nucleation, or dragon king events.

9.2 Earthquakes

The frequency-magnitude statistics of earthquakes have been recognized to satisfy log-linear scaling both globally and regionally. It is widely accepted that earthquakes satisfy the Gutenberg-Richter (GR) scaling relation (Gutenberg & Richter 1954):

$$\log N = a - bm, \tag{9.1}$$

where N is the cumulative number of earthquakes in a region and time interval with magnitudes greater than m. The scaling relation given in equation (9.1) has been shown to be equivalent to a power-law scaling between N and A_r, the earthquake rupture area (Turcotte 1997):

$$N \sim A_r^{-b}. \tag{9.2}$$

Thus earthquakes satisfy power-law scaling. Great earthquakes will be nucleation events or "characteristic earthquakes" if their magnitude significantly exceed the extrapolated scaling relation given in equation (9.1).

In this section we will consider global seismicity. Because of the many problems associated with the magnitudes of large earthquakes, the preferred approach to global seismicity is to use the Global Central Moment Tensor catalog (`www.globalcmt.org`) Dziewoński et al. (1981), Ekström et al. (2005). We utilize this catalog for the period January 1, 1977 to September 30, 2010. In order to update our results through August 15, 2011 we use moment magnitudes given in the ANSS catalog (`www.ncedc.org/cnss/`). Using

these catalogs the cumulative number of global earthquakes with magnitudes greater than for the period 1977 to August 15, 2011 is given in Figure 9.1. The roll over for small magnitudes $m_w \leq 5.25$ is attributed to the sensitivity limit of the global network used to obtain moment magnitudes m_w Engdahl et al. (1998). The roll over for $m_w \geq 7.5$ is more controversial (Rundle 1989, Sornette et al. 1996). It is usually attributed to the transition from small earthquakes with near equal lengths and depths to large earthquakes with lengths much larger than depths.

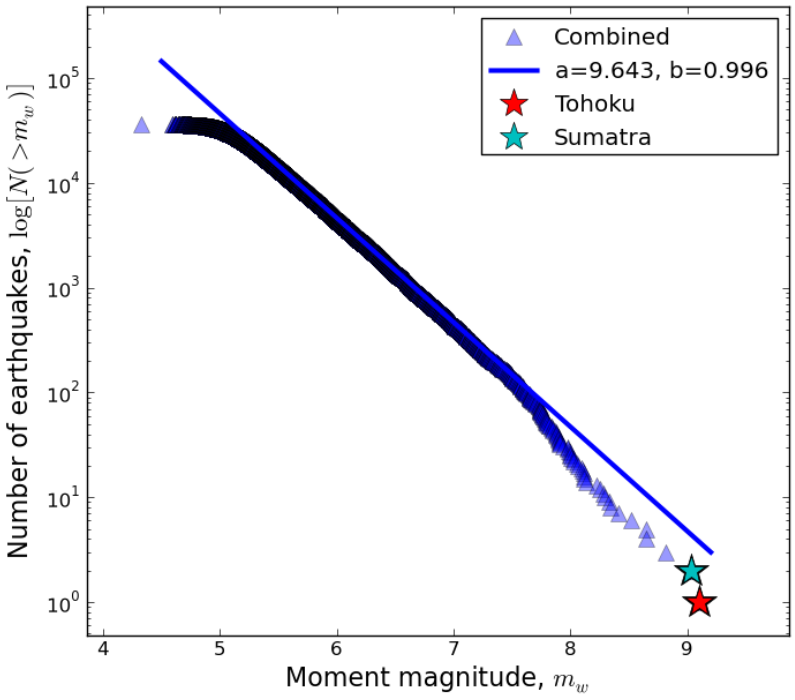


Fig. 9.1: Cumulative number of global earthquakes with magnitude greater than m_w are given as a function of moment magnitudes m_w. Observed values for the period January 1, 1977 to September 30, 2010 are obtained from the global CMT catalog, values for the period October 1, 2010 to August 15, 2011 are obtained from the ANSS catalog. The least-squares best fit of equation (9.1) to the values in the range $5.5 \leq m_w \leq 7.5$ is given taking $a = 9.643$ and $b = 0.996$. Also included are the 2004 $m_w = 9.1$ Sumatra earthquake and the 2011 $m_w = 9.1$ Tohoku earthquake.

Included in Figure 9.1 is the least squares fit of equation (9.1) to the data taking $a = 9.643$ and $b = 0.996$. The fit is carried out between $m = 5.5$ and 7.5 and includes some 30,000 earthquakes. Also included in Figure 9.1 are the $m_w = 9.1$ Sumatra earthquake on December 26, 2004 and the $m_w = 9.1$ Tohoku (Japan) earthquake on March 11, 2011. These were the largest earthquakes during the study period. The tsunami generated by the Sumatra earthquake killed some 230,000 people. The earthquake and tsunami generated by the Tohuko earthquake killed some 22,000 people. In addition the tsunami resulted in a nuclear meltdown at the Fukushima power plant (Hirose 2012). This meltdown created serious economic disruption in Japan and threatens to curtail the global use of nuclear power plants to reduce the emission of greenhouse gasses. The size and impact of the Sumatra and Tohoku earthquakes clearly qualify them as black swans. However, since they lie below the extrapolation of the power-law scaling they would not be considered nucleation events or characteristic earthquakes on a global scale. We will argue in the next section that earthquakes do exhibit characteristic behavior on a regional scale.

9.3 Characteristic Earthquakes

There are two limiting hypotheses for the behavior of faults. In the first, each fault (or fault segment) has a sequence of earthquakes that rupture the entire fault (or fault segment). The global GR (power-law, fractal) distribution of earthquakes illustrated in Figure 9.1 is attributed to a power-law (fractal) distribution of fault areas. Each fault has an earthquake with rupture area equal to the area of the fault. In seismology these are known as "characteristic" earthquakes. The other limiting hypothesis is that every fault has a GR distribution of earthquake magnitudes. The global GR scaling is the sum of the GR scaling on individual faults. The actual behavior of the earth lies between these two limits (Wesnousky 1994).

Ideally, observations would discriminate between the two limiting hypotheses. However, it is impossible to attribute smaller earthquakes to specific faults because only the largest faults can be mapped and identified. Also, location errors of smaller earthquakes make their association with each other and individual faults very difficult. The generally accepted view in seismology is that smaller earthquakes on a fault obey power-law (GR) scaling but a large fraction of the deformation on the fault is associated with large quasi-periodic "characteristic" earthquakes. Thus these large characteristic earthquakes satisfy the condition to be fault-wide nucleation events. Wesnousky (1994) has given data to support the generally accepted view that a working definition of characteristic earthquakes is that they are large earthquakes on plate-boundary faults. As examples we consider two major plate-boundary faults: *1*) The San Andreas fault in California is a major boundary fault be-

tween the Pacific and North American plates and *2*) the subduction zone fault (unnamed) above the Pacific plate as it is being subducted beneath Japan.

A comprehensive study of characteristic earthquakes has been carried out on the southern section of the San Andreas fault. Paleoseismic studies using radiocarbon dating of fluidized sediments at the Wrightwood site (Biasi et al. 2002) date characteristic earthquakes at (in years CE): 534 (407-628), 634 (551-681), 697 (657-722), 722 (695-740), 781 (736-811), 850 (800-881), 1016 (957-1056), 1116 (1047-1181), 1263 (1191-1305), 1487 (1448-1518), 1536 (1508-1569), 1685 (1647-1717), 1812 (historic), 1857 (historic). The ranges of values are the 95% confidence intervals on the radiocarbon dates. The 1857 earthquake was an historic earthquake that ruptured some 400km of the fault from central California to the Los Angeles area. No seismic recordings were available at the time of this earthquake but measurements of surface displacements indicate a magnitude $m \simeq 8.2$. In the past 75 years no earthquake with $m > 5.5$ has occurred on this fault segment. If GR scaling were valid for this fault, there should be approximately 500 $m > 5.5$ earthquakes for each $m = 8.2$ earthquake. These cannot all be in the form of aftershocks, because Bath's law (Scholz & Anderson 2002) implies that the largest aftershock should be approximately $m \simeq 7$. For this largest aftershock, GR scaling then implies that there should be approximately 10 $m > 6$ aftershocks, approximately 100 $m > 5$ aftershocks, and so on. Even accounting for aftershocks of aftershocks, it is unlikely that the aftershock sequence will have 500 $m > 5.5$ events.

Seismologists generally accept that the sequence of paleo-earthquakes listed above are characteristic earthquakes (Scholz & Anderson 2002, Wesnousky 1994). Similarly the 1906 earthquake that destroyed San Francisco is generally accepted to be a characteristic earthquake on the northern San Andreas fault. However, paleoseismic sites are not available on the rupture zone of this earthquake.

Similar paleoseismic studies have dated characteristic earthquakes on the southern Nankai Trough segment of the subduction zone beneath Japan. Characteristic earthquakes occurred in 684, 887, 1099, 1361, 1605, 1707, 1854, and 1946. The dates for these earthquakes were obtained using a variety of historic and other records (Ando 1975). The Tohoku earthquake was a characteristic earthquake on the northern section of this fault and is similar to the San Francisco earthquake in that paleoseismic data are not available (Wesnousky 1994).

The best documented sequence of characteristic earthquakes occurred on the Parkfield segment of the San Andreas fault in California. Evidence suggests that earthquakes with $m \simeq 6$ occurred in 1857, 1881, 1901, 1922, 1934, 1966, and 2004 (Bakun et al. 2005). Based on seismograms the 1922, 1934, 1966, and 2004 events were remarkably similar in magnitude. The Parkfield earthquakes are globally unique in that they are a sequence of relatively small plate-boundary "characteristic" earthquakes. Thus they occur frequently and

a complete "characteristic" earthquake cycle can be studied. In order to study the seismicity associated with the 2004 "characteristic" earthquake we consider earthquakes during the period 1972 (five years after the $m = 6.0$, June, 1966 earthquake) to 2009 (five years after the $m = 5.95$, 28 September, 2004 earthquake). To isolate seismicity associated with the characteristic Parkfield earthquake we confine our study to the region where aftershocks of the 2004 earthquake were concentrated (Shcherbakov et al. 2006).

The region is elliptical, centered at 35.9°N and -120.5°W with semi-major and semi-minor axes of 0.4° and 0.15° respectively, oriented at 137°NW. Both the aftershocks and the elliptical region are shown in Figure 9.2. It is standard practice to associate the aftershock regions with correlated seismicity to study a characteristic earthquake (Hofmann 1996, Ishibe & Shimazaki 2009, Wesnousky 1994). Another advantage of the Parkfield region is the localization of the correlated seismicity. This region clearly excludes the aftershocks of the $m = 6.5$ (2004) San Simean earthquake (lower left hand corner in Figure 9.2 and the aftershocks of the $m = 6.5$ (1983) Coalinga earthquake (just above the red elliptical area).

Parkfield is the site of the highest quality local seismic network in the world (Bakun et al. 2005). This network was constructed by the U.S. Geological Survey in the 1980's in the expectation that the next characteristic earthquake would occur. In our analysis we have used the catalog provided by the Northern California Earthquake Data Center (NCSN catalog, `http://quake.geo.berkeley.edu/ncedc/`) .

The cumulative frequency-magnitude distribution of earthquakes in the Parkfield aftershock region for the period 1972 to 2009 is given in Figure 9.3. The best-fit scaling to this distribution is given by equation (9.1) with $a = 5.65$ and $b = 1$. This is the least-squares best fit to the data in the range $2.5 \leq m \leq 4.5$. Over this range there are some 3,000 data points. The roll over for $m \leq 2.5$ is attributed to a lack of sensitivity of the network for small magnitudes (Shcherbakov et al. 2006). The roll over for $m \geq 4.5$ is attributed to the relatively small number of earthquakes: $N \sim 10$ (Shcherbakov et al. 2006).

If this scaling was applicable to the characteristic earthquake ($N = 1$) it's magnitude would have been $m = 5.65$. The $m = 5.95$ Parkfield main shock clearly lies above the extrapolation of the power-law correlation of the smaller earthquakes. An important question is whether the difference between $m = 5.65$ and $m = 5.95$ can be attributed to the statistical variability of the characteristic earthquakes. Excellent seismic records are available for the 1934, 1966, and 2004 characteristic earthquakes. The magnitudes are $m = 6.0 \pm 0.1$. In addition the seismic records of the 1934 and 1966 earthquakes are essentially identical indicating near identical points of rupture initiation and propagation pattern. The 2004 earthquake has a somewhat different rupture pattern but the rupture zone is considered to very nearly identical to the earlier earthquakes. The evidence is that these are truly characteristic earthquakes that rupture the same specified segment of the San

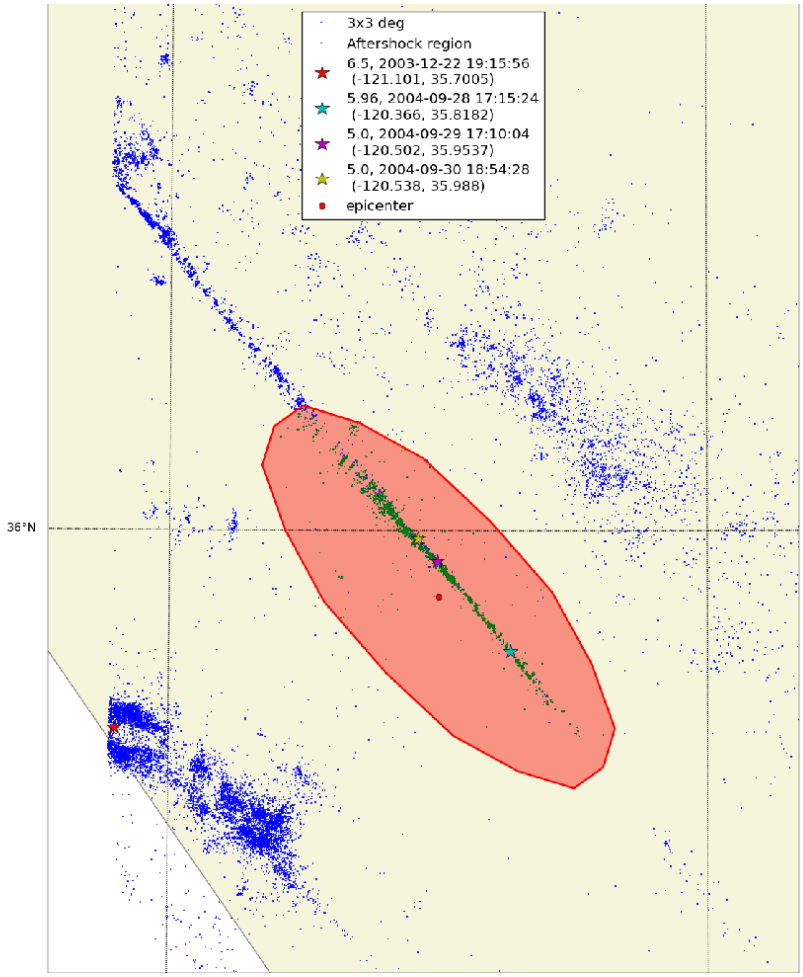


Fig. 9.2: Seismicity in central California for the period 28 September, 1999 to 28 September 2009. The aftershocks of the 18 September, 2004 Parkfield earthquake are shown in the red elliptical area. Aftershocks of the 2004 San Simean earthquake and residual aftershocks of the 1983 Coalinga earthquake are clearly seen to the south-west and north-east of the study region.

Andreas fault. The possibility that one of these characteristic earthquakes could have had a magnitude as low as 5.65 is excluded by the available data. Thus we conclude that the sequence of characteristic earthquakes of the Parkfield segment are dragon kings. Although the $\Delta m = 0.3$ difference between the 2004 Parkfield event and what one would expect from extrapolating the scaling relation in equation may not seem large, it represents about a factor

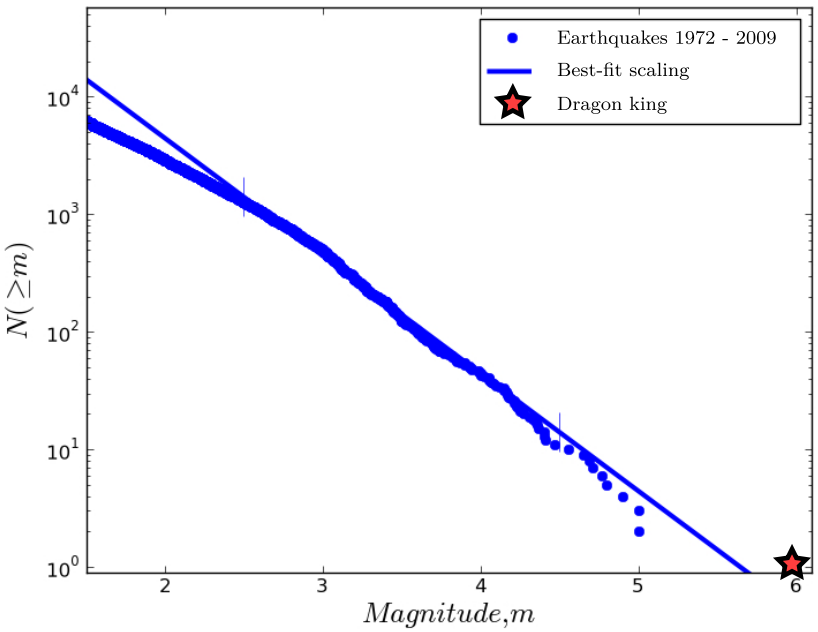


Fig. 9.3: Cumulative number of earthquakes with magnitude greater than m as a function of m for the Parkfield earthquake cycle 1972 to 2009. The best-fit scaling from equation (9.1) is also shown. The $m = 5.65$ Parkfield earthquake is shown as an outlier, sometimes called a "Dragon King".

of three difference in energies which does set it above the extrapolated GR background.

For reasons stated above the Parkfield characteristic earthquake cycle is the only such cycle that can be studied in detail using a high quality earthquake catalog. However, based on the absence of earthquakes adjacent to other segments of the San Andreas fault and other faults where characteristic earthquakes occur, we conclude that characteristic earthquakes are dragon kings with respect to the correlated seismicity (including aftershocks). The background seismicity satisfies power-law Gutenberg-Richter statistics but the characteristic earthquakes lie above the extrapolation as shown in Figure 9.3.

9.4 Models of Earthquakes

The multiple slider-block model has been proposed as a deterministic example of self-organized critical behavior (Carlson & Langer 1989). This model had previously been proposed as a simple model of earthquake behavior (Burridge & Knopoff 1967). Utilizing the multiple slider-block simulations of Abaimov et al. (2008) we will demonstrate characteristic earthquake behavior. A linear chain of 100 slider blocks was pulled over a surface at a constant velocity by a loader plate as illustrated in Figure 9.4. Each block is connected to the loader plate by a spring with spring constant k_L and adjacent blocks are connected to each other by springs with spring constant k_C. An important parameter in the problem is the ratio of spring constants $\alpha = k_C/k_L$. This is a measure of the stiffness of the system. The blocks interact with the surface through a static-dynamic friction law. With the static coefficient of friction larger than the dynamic (sliding) coefficient of friction, stick-slip behavior is observed. The size of a slip event is given by the number of blocks L that slip simultaneously in the event. The stiffness α acts as a tuning parameter in this problem.

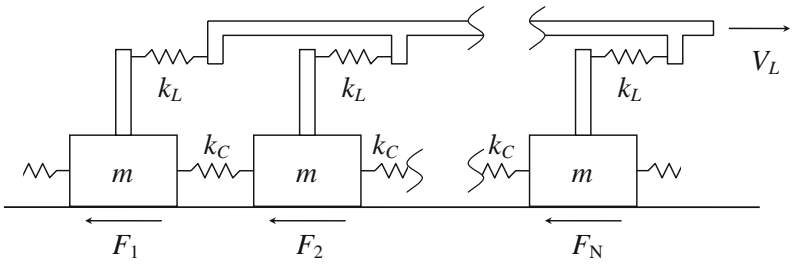


Fig. 9.4: Illustration of the one-dimensional slider-block model. A linear array of N blocks of mass m are pulled along a surface by a constant velocity V_L loader plate. The loader plate is connected to each block with a loader spring with spring constant k_L and adjacent blocks are connected by springs with spring constant k_C. The frictional resisting forces are F_1, F_2, ..., F_N.

For soft systems (α small) only small slip events occur, there is an exponential decay for larger slip events. As α is increased system wide ($L = 100$) events begin to occur. The frequency-size distribution of the slip events for a stiff system ($\alpha = 1000$) is given in Figure 9.5. Statistics for 10,000 events are given and about 1,500 are system wide ($L = 100$) events. The small events are well approximated by the power-law relation equation (9.1).

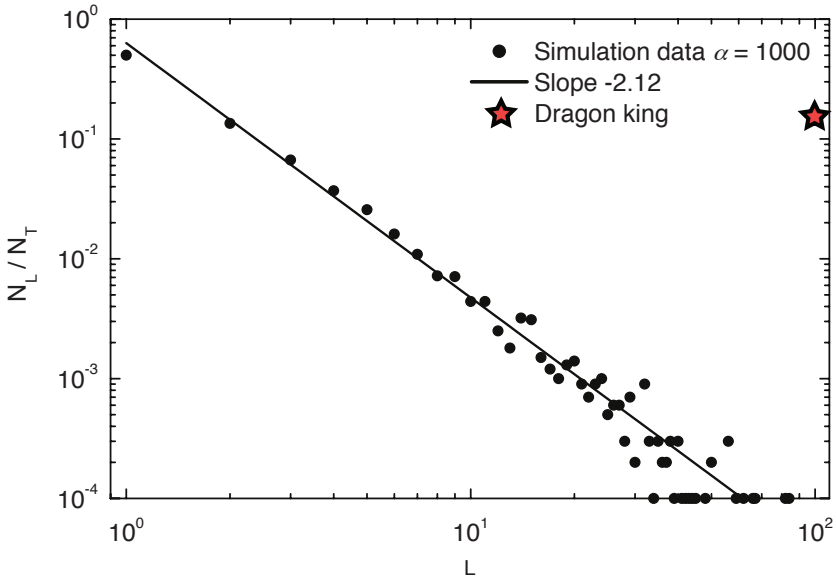


Fig. 9.5: Frequency-size density function of 10,000 slip events for a "stiff" system with $\alpha = k_C/k_L = 1000$. The ratio of the number of events N_L of event size L to the total number of events N_T is given as a function of L. The solid line is a power-law dependence with exponent $-(b+1) = -2.12$. System-wide ($L = 100$) characteristic or "Dragon-King" events are clearly evident.

9.5 Forecasting

Earthquake forecasting is of considerable importance for public safety, and to set earthquake insurance rates (Field 2007a). Recently a series of earthquake forecasts were solicited for a truly prospective forecast evaluation exercise in California (Field 2007b, Lee et al. 2011). This Relative Earthquake Likelihood Model (RELM) exercise led to the submission of nine complete forecasts by six participants for the testing region (primarily California). Participants submitted expected probabilities of occurrence of $m \geq 4.95$ earthquakes in $0.1^\circ \times 0.1^\circ$ cells for the period 1 January 1, 2006, to December 31, 2010. Probabilities were submitted for 7,682 cells in California and adjacent regions. During this period, 31 $m \geq 4.95$ earthquakes occurred in the test region. These earthquakes occurred in 22 test cells. The seismic activity was dominated by earthquakes associated with the $m = 7.2$, April 4, 2010, El Mayor-Cucapah earthquake in northern Mexico. This earthquake occurred

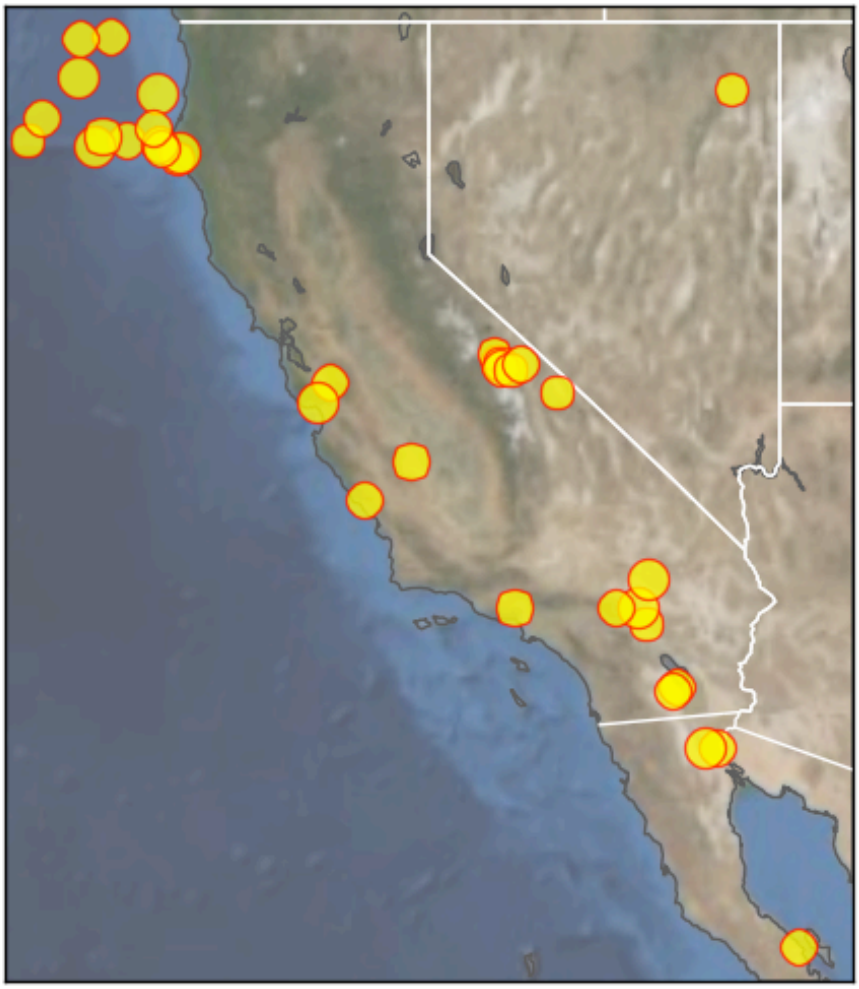


Fig. 9.6: Map of the California-Nevada region showing the epicenters of all $m_l \geq 6$ earthquakes that have occurred in the region from January 1, 1980 through July 17, 2012. Data were obtained from the ANSS catalog.

in the test region, and 16 of the other 30 earthquakes in the test region could be associated with it.

All of these forecasts relied on the Gutenberg-Richter scaling relation to partition the earthquakes into magnitude levels or "bins", and to adjust the projected rates of occurrence of the $m \geq 4.95$ events. The basic idea for most of the forecasts was to project the rate of small earthquakes, which occur at a reasonably steady rate through time, to compute the rate of the large ($m \geq 4.95$) earthquakes.

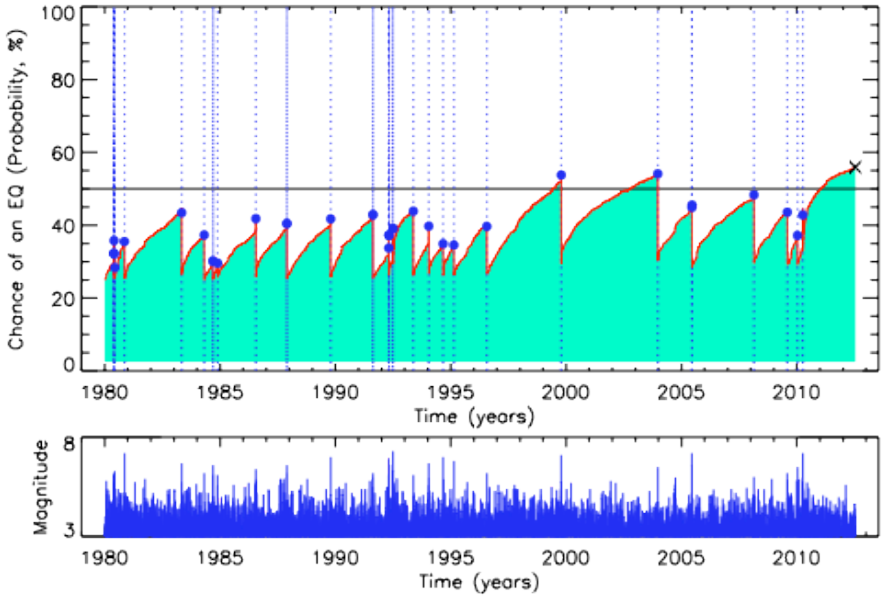


Fig. 9.7: Optimized timeseries for earthquake probabilities in the California-Nevada region. **Top**: Chance of an earthquake having magnitude $m_l \geq 6$ during the next 12 months as a function of time. **Bottom**: Magnitudes of all events in the region as a function of time. Blue dots and vertical dashed lines are earthquakes in the region having $m_l \geq 6$. "X" represents current 12-month probability of ∼56%.

Lee et al. (2011) evaluated the success of the forecasts by means of a test rewarding forecasts for the most accurate locations of the prospective earthquakes. Most of the forecasts were successful in that the great majority of the 31 earthquakes occurred in the highest probability areas. There were several "rogue" earthquakes that occurred in low probability areas, but these were very much the exception.

While this exercise in forecasting the locations of future earthquakes shows considerable promise, less progress has been made on forecasting the time of the next large earthquake in a region. Typically, the assumption is made that earthquakes occur randomly in time and should therefore be described by Poisson statistics. This leads to a probability (Yates & Goodman 2005) for the recurrence of an event during the next time interval Δt:

$$P(t) = 1 - e^{-\nu \Delta t}, \tag{9.3}$$

where ν is the long term Poisson rate of the earthquakes, and Δt is a fixed future time interval. One of the primary justifications for this assumption

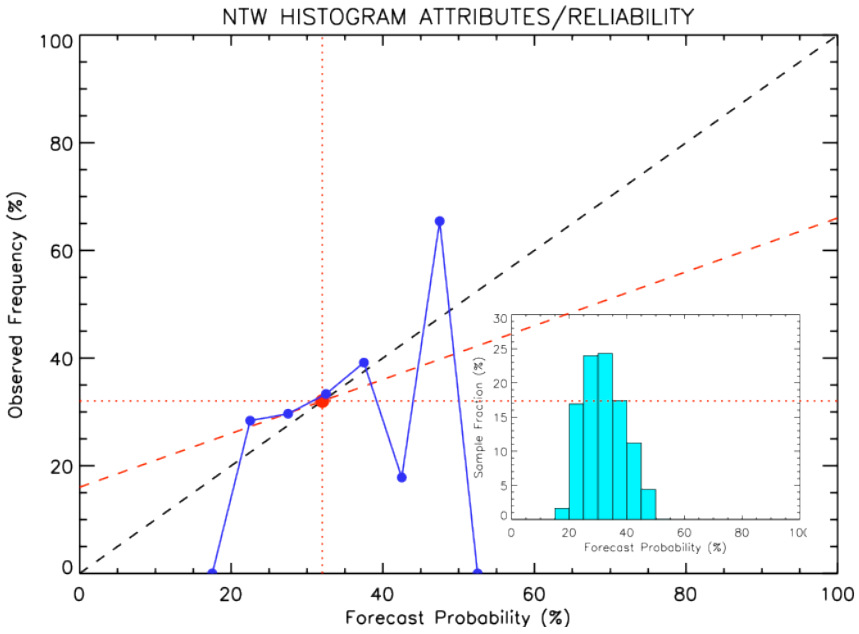


Fig. 9.8: Reliability (scatterplot) diagram for earthquake probabilities corresponding to the timeseries shown in Figure 9.7. In the figure, there are 7 bins having forecast probability, and these are plotted against observed frequencies. A perfectly reliable forecast would have zero reliability error, and the blue dots would lie along the black dashed line having slope = 1. The red dashed line is the no skill line, and the large red dot is the "climatology point", or average forecast. **Inset**: The fraction of samples in each of the 20 forecast bins of 0% to 5%, 5% to 10%, ..., 95% to 100%.

originates in a study by Gardner & Knopoff (1974), in which they removed increasing numbers of "foreshocks" and "aftershocks" from the earthquake catalog until the remaining "main shocks" could be described by a Poisson statistics. Other more recent studies assert that earthquakes should be described by negative binomial statistics (Schorlemmer et al. 2010). It should be recalled that Poisson statistics are meant to describe random events that are IID and uncorrelated such as nuclear decay, assumptions that may be questionable for earthquakes. The existence of aftershocks (e.g., Scholz (1990)) implies that earthquake correlations exist, while observations of progressively propagating mainshocks along the Anatolian fault suggest that major earthquakes are not independent (Stein et al. 1997).

As a result, we are motivated to turn to a new type of forecast, the Natural Time Weibull (NTW) model (Rundle et al. 2012). This model is based on the idea that over any period of time, the distribution of earthquake sizes should most likely be the observed Gutenberg-Richter distribution (equation (9.1)).

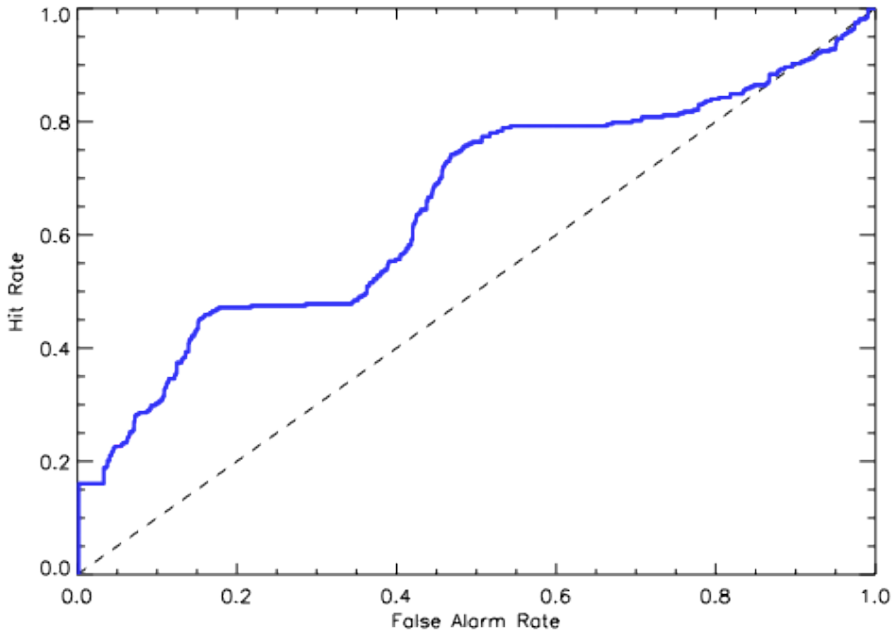


Fig. 9.9: Temporal Receiver Operating Characteristic diagram for the forecast of Figure 9.7. The solid blue line is the ROC diagram for January 1, 1980 through July 17, 2012.

For example, suppose that the large earthquake of interest has magnitude $m_l \geq 6$, and our catalog completeness level is $m_s \geq 3$. Further, suppose that the b-value in the GR relation is the typical value $b = 1$. Then for each $m_l \geq 6$ earthquake we expect 1000 $m_s \geq 3$ earthquakes. Using equation (9.1), we can compute the number of small earthquakes for every large earthquake $N_{ls} = 10^{b(m_l - m_s)}$. Another way of making the same statement is that the natural time scale for the occurrence of large earthquakes is not, in fact, the number of round trips of the earth around the sun (calendar time), but is instead the number of small earthquakes that occur between the occurrence of the large earthquakes. This can be construed as a natural time scale, or alternatively a stress-release time scale. It depends only on the observed fat-tailed statistics of the process, rather than an approximation by the Poisson law.

The NTW method is then based on the idea of counting the number of small earthquakes since the last large earthquake. Once 1000 $m_s \geq 3$ earthquakes have occurred, it is likely that another $m_l \geq 6$ earthquake is due to occur soon. To transform this idea into a probability, we use the Weibull probability law, which is often used in engineering calculations of hazard and

reliability. In the time domain, the Weibull failure law is usually stated as (Evans et al. 2000):

$$P(t) = 1 - e^{-\left(\frac{t}{\theta}\right)^{\beta}}. \tag{9.4}$$

Here t is the time to failure of the test element, θ is a time scale, and β is a constant.

Recalling the previous discussion, we apply the Weibull law (9.4) in the natural time, or event count domain. In addition, we note that earthquakes are a repetitive sequence of events. So we wish to compute the probability of observing the next large event, say $m_l \geq 6$ conditioned on the observation that n smaller events having $3 \leq m_s \leq 6$ since the last large event. In the natural time domain, this requirement leads to the equation:

$$P(\Delta n|n) = 1 - e^{-\gamma\left(\left(\frac{n+\Delta n}{N_{ls}}\right)^{\beta} - \left(\frac{n}{N_{ls}}\right)^{\beta}\right)}. \tag{9.5}$$

Here $P(\Delta n|n)$ is the conditional probability that a large earthquake $m_l \geq 6$ will occur after a number Δn of subsequent small earthquakes $3 \leq m_s \leq 6$ have occurred, given that n small earthquakes have previously occurred since the last large earthquake.

Equation (9.5) does not depend explicitly on time, or on future (calendar) time Δt. To map this probability back to the dual (calendar) time domain, we make the reasonable assumption that for a small time interval Δt:

$$\Delta n \simeq \nu_s \Delta t, \tag{9.6}$$

where ν_s is the Poisson rate of small events computed from having magnitude $3 \leq m_s \leq 6$. Using (9.6) in (9.5), we can estimate the conditional probability in the dual, time domain:

$$P(\Delta t|t) = 1 - e^{-\gamma\left(\left(\frac{n(t)+\nu_s \Delta t}{N}\right)^{\beta} - \left(\frac{n(t)}{N}\right)^{\beta}\right)}. \tag{9.7}$$

Here γ, β are constants.

We now address the question of how to determine the optimal values of γ, β. We first rewrite equation (9.7) as:

$$P(\Delta t|t) = 1 - e^{-H_R(t,\Delta t)}, \tag{9.8}$$

where $H_R(t)$ is the hazard rate (Evans et al. 2000). $H_R(t, \Delta t)$ is a function of time t since the number of small earthquakes since the last large earthquake is a function of time, $n = n(t)$.

We next use the principle that the NTW probability represents a fluctuation around the long-term Poisson rate. Thus we require that:

$$\langle H_R(t, \Delta t)\rangle_T = \nu_s \Delta t. \tag{9.9}$$

Here $\langle ... \rangle_T$ is the time average over a long (multi-decadal) time scale T considered to be much longer than the $m_l \geq 6$ recurrence interval τ:

$$\langle x(t) \rangle_T = \int_{t=-T}^{t=0} x(t')dt', \tag{9.10}$$

$t = 0$ is assumed to be the present.

Using (9.9) and (9.10) together with (9.7) and (9.8), we find that:

$$\gamma = \frac{\nu_s \Delta t}{\left\langle \left(\frac{n(t)+\nu_s \Delta t}{N} \right)^\beta - \left(\frac{n(t)}{N} \right)^\beta \right\rangle_T}. \tag{9.11}$$

Since ν_s and N are assumed to be known, Δt is a chosen time interval, and $n = n(t)$ is known from past data, γ can be determined once β is known.

To determine β, we optimize the probability using standard methods for data assimilation via backtesting with prior data. A considerable literature has accumulated in the weather and financial communities relating to forecast validation (Jolliffe & Stephenson 2011). Of particular interest are Reliability/Attributes (R/A) tests (Hsu & Murphy 1986, Murphy & Daan 1985), and Receiver Operating Characteristic (ROC) tests (Green & Swets 1966, Kharin & Zwiers 2003, Mason 1982, Murphy & Winkler 1987). The R/A test is essentially a scatter plot, in which the frequency of observed large events is plotted against the computed large event probabilities. The ROC test plots successful forecasts (H = "hit rate") against false alarms (F ="false alarm rate"). In the R/A test, reliability is computed by a weighted measure known as the Briar Skill Score (BSS), and is displayed as a scatter plot of observed frequency versus forecast probability (Hsu & Murphy 1986). In the R/A test, the BSS can be decomposed into a set of terms that represent respectively the Reliability, Sample Skill and the Resolution, which are all measures of forecast quality.

While these tests are most often used to evaluate forecasts following the forecast events, they can also be used to assimilate data into the forecast model, thereby optimizing the forecast during backtesting. It is this application that is employed to determine the value γ, β. Details are described in Rundle et al. (2012), but the basic idea is the following.

Using a grid search method, we pick a value of β, then solve (9.11) for γ. We then compute the R/A scatter plot and the ROC plot. Balancing the attributes of reliability, resolution and skill, we repeat the process many times to find an optimal value of β, and from that value, we compute γ via equation (9.11). Our basic result is that, for a broad range of regions, we find a typical optimal value of $\beta = 1.4$. Figures 9.7 - 9.9, 9.11 - 9.13 are computed using this value of β.

There are several ways to estimate uncertainty in the computation of reliability, skill and resolution. The method we use is the bootstrap method, in which large earthquake times are sampled with replacement (Rundle et al.

2011). Estimates of reliability or other parameters are computed for suites of models having the optimal parameters, and the envelope of uncertainty is computed. Details are described in Rundle et al. (2011).

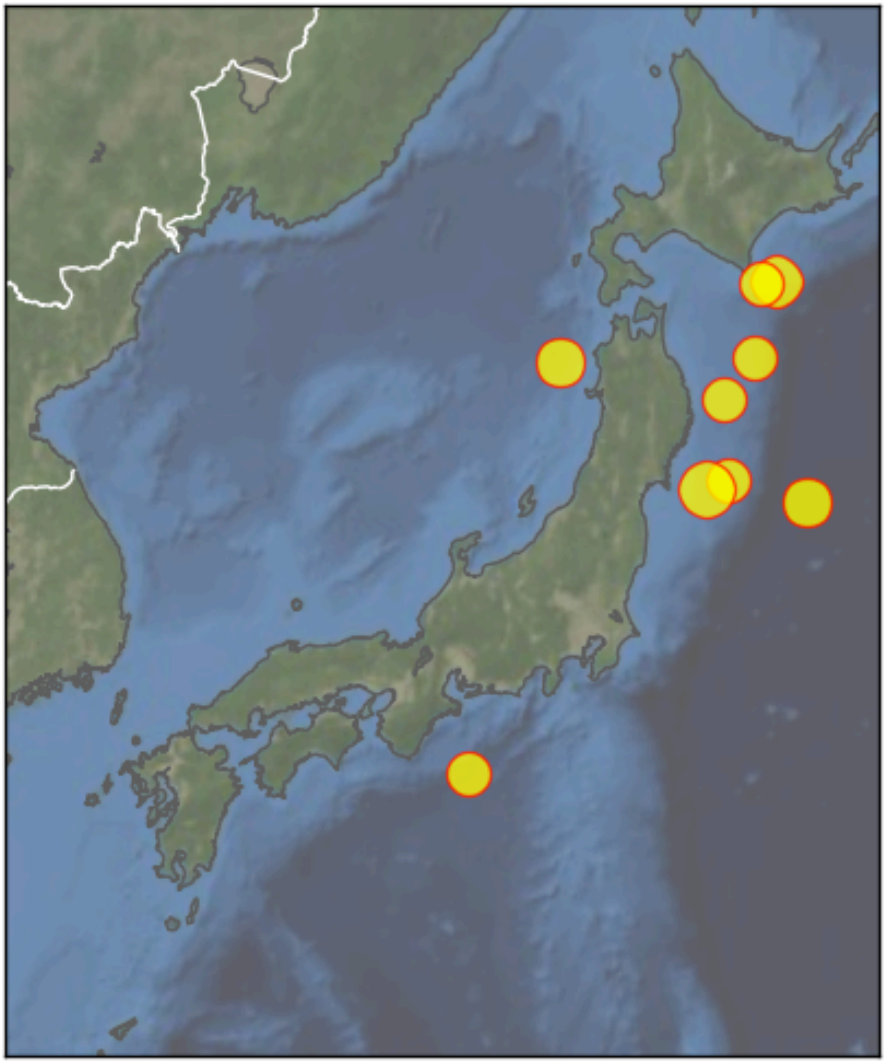


Fig. 9.10: Map of the Japan region showing the epicenters of all $m \geq 7.25$ earthquakes shallower than 40 km depth that have occurred in the region from January 1, 1980 through July 17, 2012. Data were obtained from the ANSS catalog.

In a similar way, we compute a forecast in the Japan region over the period January 1, 1980 - July 17, 2012. This forecast is for $m_l \geq 7.5$ over 4 years (48 months), in the region between latitudes 28° and 42° north, and between longitudes 127° and 146° degrees east. For the Japan calculations, we use $m_s \geq 4.25$, a maximum depth of 40 km (shallow events), and again use a time step of 0.01 year. The map of large events is shown in Figure 9.10, the forecast is shown in Figure 9.11, and the R/A and ROC curves are shown in Figures 9.12 and 9.13. In the Japan figures, the R/A and ROC curve use data only through 06:25:50.30 GMT on March 11, 2011, the date of the great $m = 9.1$ Tohoku, Japan earthquake. Again, the portion of the forecast in Figure 9.11 would be counted as a false alarm by the R/A and ROC methods if we were to use that data in calculations, so again the forecast since March 11, 2011 represents an actual forecast of future activity.

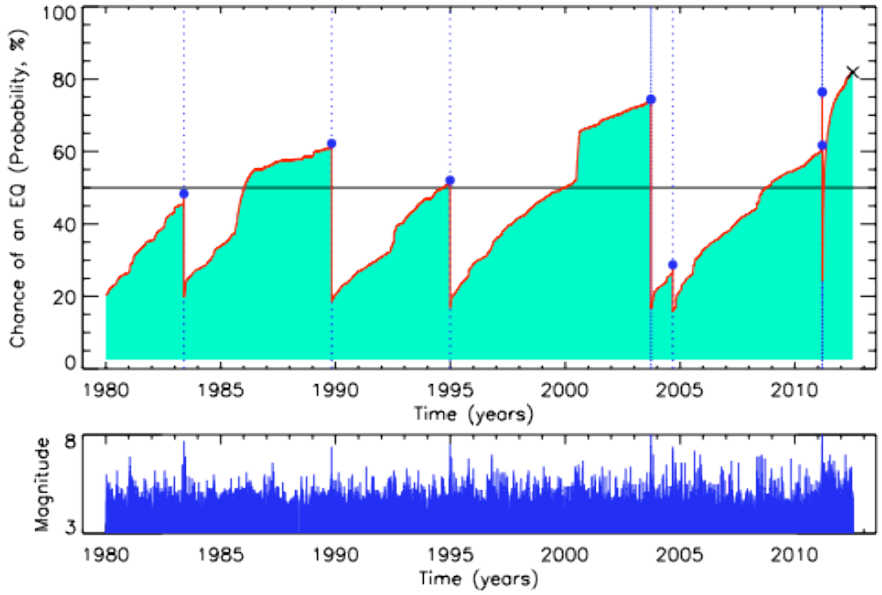


Fig. 9.11: Optimized timeseries for earthquake probabilities in Japan. **Top**: Chance of an earthquake having magnitude $m \geq 7.25$ during the next 48 months as a function of time. **Bottom**: Magnitudes of all events in the region as a function of time. Blue dots and vertical dashed lines are earthquakes in the region having $m \geq 7.25$. "X" represents current 48-month probability of ~83%.

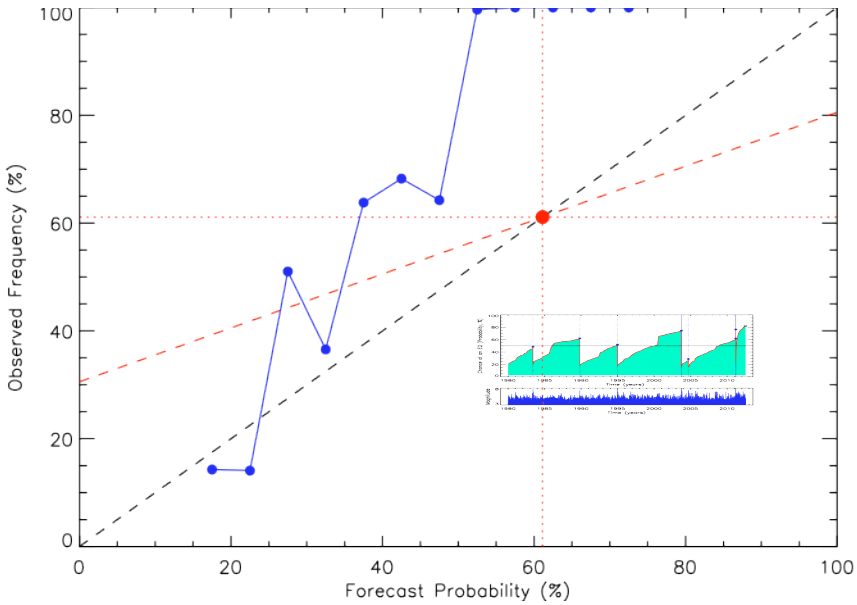


Fig. 9.12: Reliability (scatterplot) diagram for earthquake probabilities corresponding to the timeseries shown in Figure 9.11. In the figure, there are 12 bins having forecast probability, and these are plotted against observed frequencies. A perfectly reliable forecast would have zero reliability error, and the blue dots would lie along the black dashed line having slope = 1. The red dashed line is the no skill line, and the large red dot is the "climatology point", or average forecast. **Inset**: The fraction of samples in each of the 20 forecast bins of 0% to 5%, 5% to 10%, ..., 95% to 100%.

9.6 Results

Applying these ideas to earthquakes in California, we use data from the area between latitudes 29° and 42° north, and between longitudes -127° and -113° degrees west. We consider the time interval from January 1, 1980 until present (July 17, 2012), using data from the ANSS catalog with a catalog completeness level of magnitude $m_s \geq 3.5$. For data analysis and plotting purposes, we use a time step equal to 0.01 year, about 3.65 days. Larger earthquakes closer together in time than this value will not appear independently in the time series plots, although they are treated as separate events in the verification analyses. Examples include the June 28, 1992 Landers-Big Bear events ($m = 7.3$, $m = 6.5$, 3 hours apart), and the November 23-24, 1987 Elmore Ranch-Superstition Hills events ($m = 6.2$, $m = 6.6$, 12.3 hours apart).

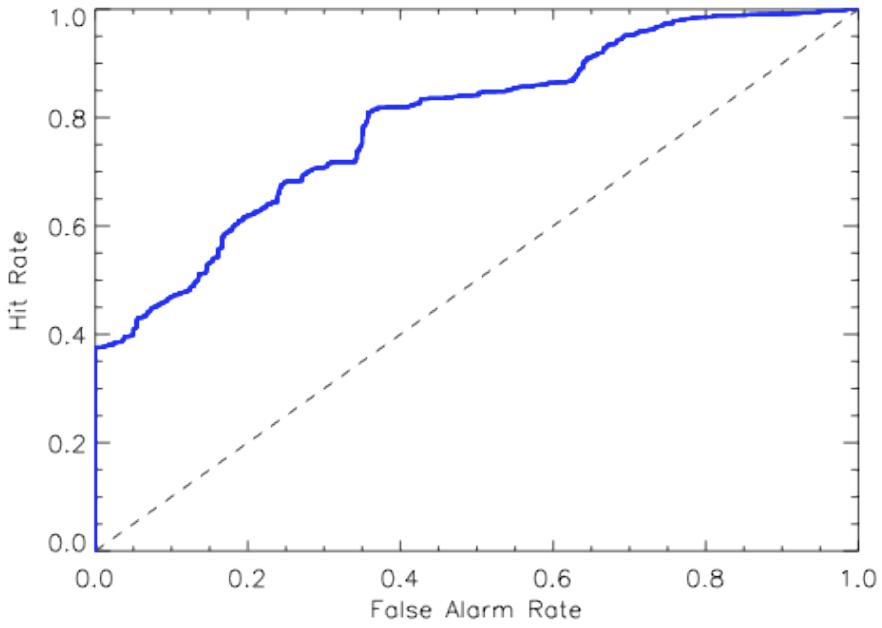


Fig. 9.13: Temporal Receiver Operating Characteristic diagram for the forecast of Figure 9.11. The solid blue line is the ROC diagram for January 1, 1980 through July 17, 2012.

To make practical calculations, we first compute the Poisson rate of small earthquakes $\nu_s(m \geq 3.5)$ in the defined region. Figure 9.6 shows a map of the region, and Figure 9.7 shows the optimal forecast as a function of time in years over the period January 1, 1980 - July 17, 2012. Figure 9.8 is a plot of the Reliability Diagram, and Figure 9.9 is a plot of the Receiver Operating Characteristic. Details are given in Rundle et al. (2012, 2011).

For the backtesting process, we use data only up until 22:40:41.77 GMT on April 4, 2010, the time and date of the $m = 7.2$ El Mayor-Cucapah (Baja) earthquake. The reason is that any forecast arising from any of the data from April 4 until present would automatically be counted as false alarm, since by definition, no earthquake with $m_l \geq 6$ has occurred since April 4, 2010. Therefore, any forecast after April 4, 2010, for example in Figure 9.7, can be considered as an actual future forecast for $m_l \geq 6$ earthquakes within the California-Nevada study area.

In a similar way, we compute a forecast in the Japan region over the period January 1, 1980 - July 17, 2012. This forecast is for $m_l \geq 7.5$ over 4 years (48 months), in the region between latitudes 28° and 42° north, and between longitudes 127° and 146° degrees east. For the Japan calculations, we use $m_s \geq 4.25$, a maximum depth of 40 km (shallow events), and again use a

time step of 0.01 year. The map of large events is shown in Figure 9.10, the forecast is shown in Figure 9.11, and the R/A and ROC curves are shown in Figures 9.12 and 9.13. In the Japan figures, the R/A and ROC curve use data only through 06:25:50.30 GMT on March 11, 2011, the date of the great $m = 9.1$ Tohoku, Japan earthquake. Again, the portion of the forecast in Figure 9.11 would be counted as a false alarm by the R/A and ROC methods if we were to use that data in calculations, so again the forecast since March 11, 2011 represents an actual forecast of future activity.

9.7 Summary

We have discussed the role of scaling and self-organization in both earthquake data and earthquake models. We have also discussed the problem of earthquake forecasting, and show that forecasts can be developed based on the idea of "filling in" a fat-tailed (scaling) distribution. We then show that earthquake forecasts can be computed, with validation and verification based on standard methods in the literature.

As of this writing (July 17, 2012), it can be seen from Figures 9.7 and 9.11 that significant large earthquake activity is expected in the relatively near future in both California-Nevada ($m_l \geq 6$, $\Delta t \leq 1$ year from 2012/7/17) and Japan ($m_l \geq 7.25$, $\Delta t \leq 4$ years from 2012/7/17).

Acknowledgements

Research by MKS was supported by NASA Fellowship NNX11AL92H. Research by JBR and JRH was performed with funding from NASA grant NNX08AF69G to the University of California, Davis.

References

Abaimov, S., Turcotte, D. L., Shcherbakov, R., Rundle, J. B., Yakovlev, G., Goltz, C., & Newman, W. (2008). Earthquakes: Recurrence and interoccurrence times. *Pure and Applied Geophysics*, *165*, 777–795. 10.1007/s00024-008-0331-y.

Alava, M., Nukala, P., & Zapperi, S. (2006). Statistical models of fracture. *Advances in Physics*, *55*(3), 349 – 476.

Ando, M. (1975). Source mechanisms and tectonic significance of historical earthquakes along the nankai trough, japan. *Tectonophys.*, *27*, 119–140.

Bakun, W. H., Aagaard, B., Dost, B., Ellsworth, W. L., Hardebeck, J. L., Harris, R. A., Ji, C., Johnston, M. J. S., Langbein, J., Lienkaemper, J. J., Michael, A. J., Murray, J. R., Nadeau, R. M., Reasenberg, P. A., Reichle, M. S., Roeloffs, E. A.,

Shakal, A., Simpson, R. W., & Waldhauser, F. (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature*, *437*, 969–974.

Biasi, G. P., Weldon, R. J., Fumal, T. E., & Seitz, G. G. (2002). Paleoseismic event dating and the conditional probability of large earthquakes on the southern san andreas fault, california. *Bull. Seis. Soc. Am.*, *92*(7), 2761–2781.

Burridge, R., & Knopoff, L. (1967). Model and theoretical seismicity. *Bull. Seis. Soc. Am.*, *57*(3), 341–371.

Carlson, J. M., & Langer, J. S. (1989). Mechanical model of an earthquake fault. *Phys. Rev. A*, *40*(11), 6470–6484.

Dziewoński, A. M., Chou, T. A., & Woodhouse, J. W. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J. Geophys. Res.*, *86*(B4), 2825–2852.

Ekström, G., Dziewoński, A. M., Maternovskaya, N. N., & Nettles, M. (2005). Global seismicity of 2003: centroid–moment-tensor solutions for 1087 earthquakes. *Phys. Earth Planet. Inter*, *148*, 327–351.

Engdahl, E. R., van der Hilst, R., & Buland, R. (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bull. Seis. Soc. Am.*, *88*(3), 722–743.

Evans, M., Hastings, N. A. J., & Peacock, J. B. (2000). *Statistical distributions*. Wiley series in probability and statistics. Probability and statistics section. New York: Wiley, 3rd ed. ed.
URL `http://www.loc.gov/catdir/bios/wiley042/99088655.html`

Field, E. H. (2007a). Overview of the working group for the development of regional earthquake likelihood models (RELM). *Seis. Res. Lett.*, *78*, 7–16.

Field, E. H. (2007b). A summary of previous working groups on California earthquake probabilities. *Seis. Soc. Am. Bull.*, *97*, 1033–1053.

Gardner, J. K., & Knopoff, L. (1974). Is the sequence of earthquakes in southern california, with aftershocks removed, poissonian? *Bull. Seis. Soc. Am.*, *64*, 1363–1367.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Huntington, N.Y.: R. E. Krieger Pub. Co.

Guarino, A., Garcimartín, A., & Ciliberto, S. (1998). An experimental test of the critical behaviour of fracture precursors. *The European Physical Journal B - Condensed Matter and Complex Systems*, *6*, 13–24. 10.1007/s100510050521.

Gutenberg, B., & Richter, C. F. (1954). *Seismicity of the Earth and Associated Phenomena*. Princeton, N.J.: Princeton University Press.

Hirose, K. (2012). 2011 fukushima dai-ichi nuclear power plant accident: summary of regional radioactive deposition monitoring results. *J. Env. Radioactivity*, *111*, 13–17.

Hofmann, R. B. (1996). Individual faults can't produce a gutenberg-richter earthquake recurrence. *Engineering Geology*, *43*(1), 5 – 9.

Hsu, W., & Murphy, A. H. (1986). The attributes diagram a geometrical framework for assessing the quality of probability forecasts. *International Journal of Forecasting*, *2*(3), 285 – 293.

Ishibe, T., & Shimazaki, K. (2009). Seismicity in source regions of large interplate earthquakes around japan and the characteristic earthquake model. *Earth Planets and Space*, *61*(9), 1041 – 1052.

Jolliffe, I. T., & Stephenson, D. B. (2011). *Forecast verification: a practitioner's guide in atmospheric science*. Hoboken, NJ: John Wiley and Sons, 2nd ed ed.

Kharin, V. V., & Zwiers, F. W. (2003). On the roc score of probability forecasts. *Journal of Climate*, *16*(24), 4145–4150.
URL `http://dx.doi.org/10.1175/1520-0442(2003)016<4145:OTRSOP>2.0.CO;2`

Kinnison, R. R. (1985). *Applied Extreme Value Statistics*. Columbus, Ohio: Battelle Press.

Lee, Y. T., Turcotte, D. L., Holliday, J. R., Sachs, M. K., Rundle, J. B., Chen, C. C., & Tiampo, K. F. (2011). Results of the regional earthquake likelihood models (RELM) test of earthquake forecasts in California. *Proc. Nat. Acad. Sci. (USA)*, *108*(40), 16533–16538.

Mantegna, R. N., & Stanley, H. E. (2000). *An introduction to econophysics: correlations and complexity in finance*. Cambridge, UK: Cambridge University Press. URL `http://www.loc.gov/catdir/samples/cam034/99028047.html`

Mason, I. B. (1982). A model for assessments of weather forecasts. *Austral. Met. Mag.*, *30*, 291–303.

Murphy, A. H., & Daan, H. (1985). Forecast evaluation. In A. H. Murphy, & R. W. Katz (Eds.) *Probability, Statistics and Decision Making in the Atmospheric Sciences*. Westview Press Boulder, CO.

Murphy, A. H., & Winkler, R. L. (1987). A general framework for forecast verification. *Mon. Weather Rev.*, *115*, 1330–1338.

Rundle, J. B. (1989). Derivation of the complete gutenberg-richter magnitude-frequency relation using the principle of scale invariance. *J. Geophys. Res.*, *94*(B9), 12337–12342.

Rundle, J. B., Holliday, J. R., Graves, W., Turcotte, D. L., Tiampo, K. F., & Klein, W. (2012). Probabilities for large events in driven threshold systems. *Phys. Rev. E*, *86*(021106).

Rundle, J. B., Holliday, J. R., Yoder, M., Sachs, M. K., Donnellan, A., Turcotte, D. L., Tiampo, K. F., Klein, W., & Kellogg, L. H. (2011). Earthquake precursors: activation or quiescence? *Geophysical Journal International*, *187*(1), 225–236.

Sachs, M. K., Yoder, M. R., Turcotte, D. L., Rundle, J. B., & Malamud, B. D. (2012). Black swans, power laws, and dragon-kings: Earthquakes, volcanic eruptions, landslides, wildfires, floods, and soc models. *Eur. Phys. J. Special Topics*, *205*, 167–182.

Scholz, C. H. (1990). Geophysics–earthquakes as chaos. *Nature*, *348*, 197–198.

Scholz, C. H., & Anderson, D. L. (2002). *The Mechanics of Earthquakes & Faulting*. Cambridge: Cambridg University Press, 2nd ed.

Schorlemmer, D., Zechar, J. D., Werner, M. J., Field, E. H., Jackson, D. D., & Jordan, T. H. (2010). First results of the regional earthquake likelihood models experiment. *Pure Ap. Geophys.*, *167*, 859–876.

Shcherbakov, R., Turcotte, D. L., & Rundle, J. B. (2006). Scaling properties of the parkfield aftershock sequence. *Bull. Seis. Soc. Am.*, *96*(4B), S376–384.

Sornette, D. (2004). *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press.

Sornette, D. (2009). Dragon-kings, black swans and the prediction of crises. *International Journal of Terraspace Science and Engineering*, *2*, 1.

Sornette, D., Knopoff, L., Kagen, Y., & Vanneste, C. (1996). Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes. *J. Geophys. Res.*, *101*(B6), 13883–13893.

Stein, R. S., Barka, A. A., & Dieterich, J. H. (1997). Progressive failure on the north anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International*, *128*(3), 594–604. URL `http://dx.doi.org/10.1111/j.1365-246X.1997.tb05321.x`

Taleb, N. (2007). *The Black Swan: The Impact of the Highly Improbable*. New York: Random House, 1st ed ed.

Turcotte, D. L. (1997). *Fractals and Chaos in Geology and Geophysics*. Cambridge, U.K.: Cambridge University Press, 2nd ed ed.

Wesnousky, S. G. (1994). The gutenberg-richter or characteristic earthquake distribution, which is it? *Bull. Seis. Soc. Am.*, *84*(6), 1940–1959.

Yates, R. D., & Goodman, D. J. (2005). *Probability and Stochastic Processes*. New York: John Wiley and Sons.

Chapter 10
Wildfires and the Forest-Fire Model

Stefan Hergarten

The Drossel-Schwabl forest-fire model (Drossel and Schwabl 1992), often referred to as *the* forest-fire model, is one of the three most widely studied models in the context of SOC. Remarkably, it is in principle identical to a self-organized percolation model proposed some years earlier by Henley (1989). Beyond this, it is an impressive example of modeling real-world phenomena since its relevance to nature was discovered several years after the model was proposed. In contrast, the first and most widespread model of SOC, the Bak-Tang-Wiesenfeld model (Bak et al 1987, 1988), is often entitled sandpile model, and the model results were compared with real sandpiles and ricepiles (Frette et al 1996), but its relationship to sandpile dynamics is still vague (Hergarten 2002). So the Bak-Tang-Wiesenfeld model is still some kind of paradigm of a self-organized critical system, but it is more an abstract representation of general avalanching phenomena than a model for any real-world phenomenon. In contrast, the third of the tree most widespread models of SOC, the Olami-Feder-Christensen model (Olami et al 1992) aimed at explaining known scale-invariant properties of earthquakes from the beginning. The forest-fire model was believed to be too much oversimplified to capture any important properties of wildfire dynamics until Malamud et al (1998) put real wildfires into the context of SOC using wildfire size distributions from different regions. Although the model is indeed too simple to describe wildfire dynamics in detail, there is growing evidence that it even captures more properties of wildfires than just the size distribution (Zinck and Grimm 2008). Furthermore, differences between natural and human-induced forest fires have been reproduced correctly (Krenn and Hergarten 2009). However, its acceptance in the scientific community of fire ecology is still not very high.

Stefan Hergarten
Institut für Geo- und Umweltnaturwissenschaften, Albert-Ludwigs-Universität Freiburg, e-mail: stefan.hergarten@geologie.uni-freiburg.de

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

10.1 The Forest-Fire Model

The Drossel-Schwabl forest-fire model is a stochastic cellular automaton model that is usually considered on a two-dimensional square lattice with $L \times L$ sites and periodic boundary conditions.

The forest-fire model published by Bak et al (1990) may be seen as the ancestor of the Drossel-Schwabl forest-fire model. In this model, each site can be either empty or occupied by a tree which may be green or burning. In each step, the lattice is updated according to the following rules:

1. A green tree catches fire if any of its four nearest neighbors is burning.
2. A burning tree turns into an empty site.
3. At each empty site, a green tree grows with a given probability p.

These rules are applied simultaneously to all sites.

The results presented in the original paper pointed towards SOC. At least under certain conditions, the model evolves towards a quasi-steady state where the number of burning trees in each step is power-law distributed. However, it was soon recognized that this model suffers from some problems. In principle, it does not simulate sequences of fires, but only one fire which extends over the whole simulation and is kept alive by the regrowth of trees. Apart from the problem that the fire might die, this behavior is not very realistic with respect to forest fires in reality. Further simulations (Grassberger and Kantz 1991, Moßner et al 1992) revealed that this forest-fire model is not SOC, but exhibits a mainly regular behavior with spiral-shaped fire fronts.

Drossel and Schwabl (1992) fixed the shortcoming that there is just one fire by introducing a spontaneous ignition of green trees as it takes place in reality by lightning or by human impact by adding the following rule:

4. A green tree becomes a burning tree with a probability f even if none of its neighbors is burning.

In addition, they introduced a separation of time scales. In analogy to real forest fires, burning down a cluster of trees takes place much more rapidly than raising new trees. Therefore, it is assumed that growth and spontaneous ignition stop as long as any trees are burning. Formally, this separation of time scales can be achieved by performing the limit $p \to 0$ and $f \to 0$, while the ratio $\theta = \frac{p}{f}$ is kept constant. In a computer model, the separation of time scales can be realized by burning down the entire cluster of trees connected to a tree that has been ignited within one model step. This leads to a modified set of rules for updating the lattice in each step:

1. Each tree is ignited with a probability f. The cluster of trees connected to an ignited tree is burnt down immediately, i.e., the corresponding sites become empty.
2. At each empty site, a green tree grows with a probability p.

As a consequence of the separation of time scales, only two states are required instead of three, there is no need to consider burning sites explicitly. Thus, the model can be implemented with a storage requirement of only one bit per site, which is the minimum storage requirement for a model where each site has at least one degree of freedom. This allows larger lattices than all other models in the field of SOC.

This version of the model suggested by Drossel and Schwabl (1992) is essentially the same as the self-organized percolation model published by Henley (1989) three years before. Maybe the detour to the version of Bak et al (1990) was necessary to put this model into the wider context of SOC, so that the model is today mostly associated with the names Drossel and Schwabl instead of Henley.

The two rules given above come quite close to the simplest ideas on the propagation of forest fires in reality, but the numerical realization is cumbersome. In each step, all sites must be checked for either the growth of a new tree or for ignition. Therefore, the numerical effort per step is proportional to the total number of sites, even if not any site is finally ignited. Therefore, a further modification was suggested (Grassberger 1993, Clar et al 1994) in order to make the model feasible on large grids. Based on the easily recognized property that $\theta = \frac{p}{f}$ attempts to plant new trees are made between two ignition events in the mean, it is assumed that *exactly* θ attempts are made until a spark is thrown. So the model rules turn into what is usually referred to as *the* forest-fire model:

1. A randomly chosen site is ignited. If it is occupied by a tree, this tree and all trees connected by nearest-neighbor relations to it are immediately burnt down.
2. A total of θ new trees is randomly placed on the grid. If a site is already occupied by a tree, the new tree is ignored.

Unfortunately, the use of the symbol θ is not unique in the literature. Several authors (e.g., Grassberger 1993, 2002, Pastor-Satorras and Vespignani 2000, Krenn and Hergarten 2009, Hergarten and Krenn 2011) use it as it is defined above and several authors retain $\frac{p}{f}$ or $\frac{f}{p}$ (e.g., Drossel and Schwabl 1992, Christensen et al 1993, Clar et al 1994, Schenk et al 2002). On the other hand, some authors introduce θ just the other way round, $\theta = \frac{f}{p}$ (Vespignani and Zapperi 1998, Pruessner and Jensen 2002, 2004), so that their parameter θ is the inverse of θ defined above.

10.2 Numerical and Theoretical Results

In the last 20 years, considerable effort has been spent to understand the behavior of the forest-fire model.

It was, of course, immediately recognized that the forest-fire model self-organizes towards a quasi-steady state with roughly power-law distributed events up to a cutoff size that depends on the growth rate θ. In this state, the mean tree density ρ, i.e., the number of occupied sites divided by the total number of sites, is about 0.41.

The dependence of the large-size cutoff on θ can be easily understood by balancing the number of trees destroyed by fires with the number of growing trees in equilibrium. Since trying to plant a tree at a site that is already occupied has no effect, $(1-\rho)\theta$ new trees grow in each step. If $\overline{s}$ is the mean fire size, the mean number of trees destroyed per model step is $\rho\overline{s}$, so that equilibrium requires

$$\overline{s} = \frac{1-\rho}{\rho}\theta \approx 1.44\theta. \tag{10.1}$$

Thus, criticality can only be expected in the limit $\theta \to \infty$. As the total size of the lattice should be much larger than the largest fires, this requires $L \to \infty$, too. But since the model only requires one bit per site, large lattices are not such a problem here.

In contrast to, e.g., the Bak-Tang-Wiesenfeld model, the forest-fire model generates quite patchy spatial patterns consisting of regions with different tree densities. Figure 10.1 provides an example. On a qualitative level, this behavior can be easily understood. A large fire clears an area almost completely, while very few small patches of trees survive. Thus, the tree density is almost zero in a region just burnt down. During regrowth of such an area, large fires are unlikely since they require large clusters of trees and thus a high tree density. As the trees are randomly distributed in the domain, the probability of large fires remains low until the density approaches the threshold of site percolation for the square lattice (e.g., Stauffer and Aharony 1994), $\rho_p = 0.59275$ (Newman and Ziff 2000). Thus, the density oscillates more or less periodically like a saw-tooth function between almost zero and ρ_p even on large scales. As a result, Fig. 10.1 shows large areas which were recently burnt as well as large areas of high density which are prone to large fires in future.

However, quantitative understanding of this behavior is still poor. The mean density ρ has been addressed in several numerical studies, resulting in an extrapolation $\rho \to 0.4084 \pm 0.0005$ for $\theta \to \infty$ (Pastor-Satorras and Vespignani 2000). But there is still no theory how the mean density relates to the temporal variability in the density on all scales up to the largest events, and whether it is just a coincidence that the mean density is close to the density of the empty sites in site percolation, i.e., $1-\rho_p$ (Pruessner and Jensen 2002). So it seems that even better estimates of ρ may be of limited value, and furthermore it seems to be unlikely that the exact value of ρ has any relevance to modeling real-world phenomena.

The scaling exponent τ of the event-size distribution

$$f(s) \propto s^{-\tau} \tag{10.2}$$

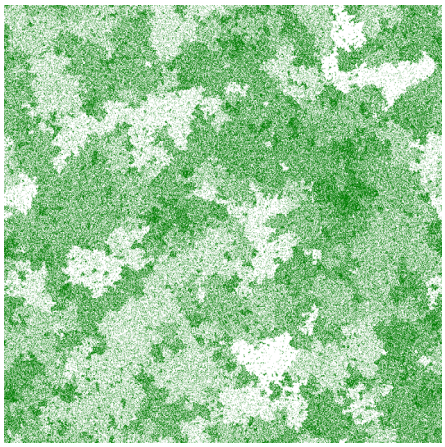


Fig. 10.1: A snapshot from a simulation of the forest-fire model with $\theta = 1024$ on a 1024×1024 lattice.

is, of course, the most important property of a self-organized critical system. Apart from the theoretical interest, it is also the most manifest property to be compared with real-world data. Drossel and Schwabl (1992) suggested $\tau = 1$ first, but this was soon refuted and replaced with estimates $\tau \in [1.14, 1.16]$ (Christensen et al 1993, Grassberger 1993, Henley 1993, Clar et al 1994, Honecker and Peschel 1997). It may be surprising that the variability in the estimates of τ increased in the following years. Pastor-Satorras and Vespignani (2000) obtained $\tau = 1.08$ by analyzing the moments of the distribution instead of the slope in a double-logarithmic plot. Schenk et al (2002) even suggested $\tau = 1.43$, but this value was partially based on theoretical arguments which were refuted. Hergarten (2002) obtained $\tau = 1.23$, and the presumably most widely accepted value is $\tau = 1.19$ (Grassberger 2002, Pruessner and Jensen 2002, 2004, Hergarten and Krenn 2011).

Figure 10.2 illustrates why it is difficult to give a reliable estimate of τ in the limit $\theta \to \infty$. The breakdown of the power law at large event sizes is accompanied by an excess of fires within a certain range of sizes, visible as a bump in the distribution. This bump becomes even relatively larger if θ increases, leading Grassberger (2002) to the conclusion that the forest-fire model violates the simple scaling behavior found, e.g., in the Bak-Tang-Wiesenfeld model. Grassberger (2002) found a different scaling exponent $\tau = 1.11$ if an envelope of the curves including the bump is considered instead of the apparently straight part left of the bump, so that the question may be raised whether it makes sense to define a scaling exponent τ at all for the forest-fire model (Pruessner and Jensen 2004).

In the childhood of the forest-fire model in the 1990s, numerical facilities were far from being able to address these questions adequately. For

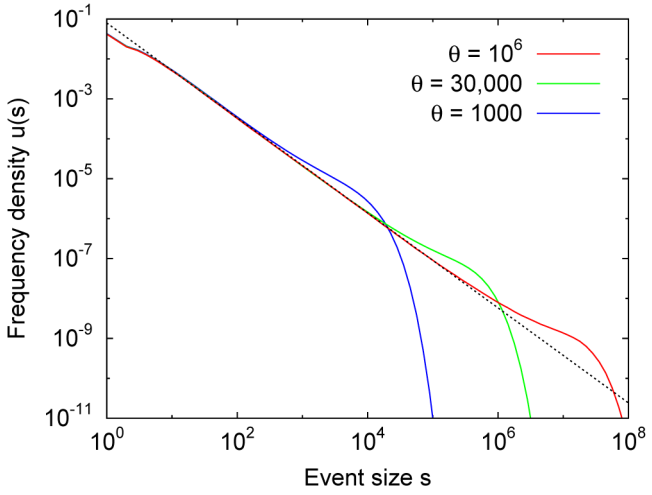


Fig. 10.2: Non-cumulative frequency density of the fires in the forest-fire model for different growth rates θ (Hergarten and Krenn 2011). The dashed line corresponds to a power law with an exponent $\tau = 1.19$.

instance, the large-scale simulations of Honecker and Peschel (1997) involved only growth rates up to $\theta = 10,000$.

A theoretical approach to determine the scaling exponent τ was recently published by Hergarten and Krenn (2011). This approach still does not predict τ from the model rules alone, but relates it to a geometrical property of the clusters of trees, the so-called accessible perimeter. In this approach, clusters are characterized only by their size without regard to spatial correlations between clusters. In its spirit, the idea is similar to the hierarchical clustering idea of Gabrielov et al (1999) who obtained $\tau = 1$ under some extreme simplifications.

In the forest-fire model, each new tree planted on the grid either forms a new cluster of size one, extends the size of an existing cluster by one or connects two or more existing clusters. Growth of clusters by individual new trees depends on the total perimeter p_t of the clusters, consisting of all empty sites adjacent to sites in the cluster. Since clusters are not dense in general, the total perimeter includes both external and internal empty nearest-neighbor sites and increases almost linearly with the cluster size s in the mean:

$$p_t(s) = fs \tag{10.3}$$

with $f \approx 0.7$. Coalescence of existing clusters is obviously the most efficient process to make mid-sized and large clusters even larger, but also more complicated than growth by individual new trees. It was conjectured that coalescence does not depend on the total perimeter of clusters, but on the accessi-

ble perimeter p_a consisting of those perimeter sites which can be reached, in principle, by a random walker coming from infinity (Grossman and Aharony 1986). The difference between accessible and total perimeter is illustrated in Fig. 10.3. It was numerically found that the accessible perimeter has a fractal scaling relation

$$p_a(s) = gs^h \tag{10.4}$$

with $g = 3$ and $h = 0.69$ for intermediate cluster sizes. As discussed by Krenn and Hergarten (2009), this behavior is almost identical to that obtained for site percolation clusters by combining numerical results on the fractal dimension of the accessible perimeter (Grossman and Aharony 1986) with theoretical arguments.

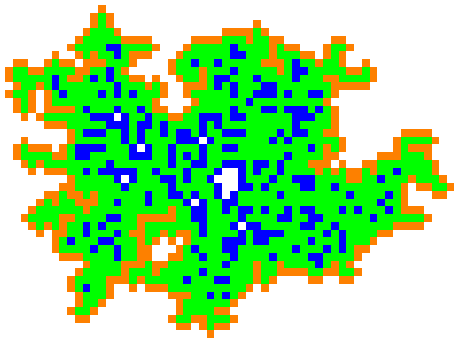


Fig. 10.3: Illustration of the total and the accessible perimeter of a cluster. Sites belonging to the accessible perimeter are orange. Internal perimeter sites, i.e., sites belonging to the total perimeter, but not to the accessible perimeter, are colored blue.

Neglecting higher-order coalescence (more than two clusters) and assuming that the process is most efficient if both clusters are of similar sizes, Hergarten and Krenn (2011) derived the ordinary differential equation

$$\frac{1}{2s}\frac{d}{ds}\left(sp_a(s)^2u(s)^2\right) + f\frac{d}{ds}u(s) + \frac{u(s)}{\theta} = 0. \tag{10.5}$$

The first term refers to coalescence, the second to growth by individual trees, and the third to the loss by burning. It was shown that the first term dominates for mid-sized and large clusters, leading to

$$sp_a(s)^2u(s)^2 = \text{const}, \tag{10.6}$$

and thus in combination with Eq. (10.4)

$$u(s) \propto s^{-(h+\frac{1}{2})}. \tag{10.7}$$

So the frequency density of the fires follows a power law with an exponent

$$\tau = h + \frac{1}{2} = 1.19 \tag{10.8}$$

in perfect agreement with the presumably most reliable numerical estimates (Grassberger 2002, Pruessner and Jensen 2002, 2004).

Hergarten and Krenn (2011) also suggested an explanation for the bump in the distribution at large event sizes. The idea hinges on the finding that very large clusters have less smooth boundaries than smaller clusters. This was attributed to the dominance of coalescence over growth by individual trees which smoothes the boundary. Although reasonable, this approach requires some ad hoc assumptions, and the exponential decay at the right-hand side of the bump of the distribution is not predicted well.

After all, knowledge on the forest-fire model is still far off from being a complete theory. However, at least the presumably most important property of the model, the scaling exponent $\tau = 1.19$, can be derived from known scaling properties of percolation clusters. In return, knowledge on the mean tree density ρ going beyond numerical results is still sparse.

10.3 The Relationship to Real Wildfires

At least until the seminal paper published by Malamud et al (1998), all publications on the forest-fire model only addressed the model's behavior in itself. It seems that nobody dared to think about a potential relevance of such an oversimplified model to real wildfires. As mentioned in the introduction of this chapter, it was much easier for the Olami-Feder-Christensen earthquake model as it was immediately found to reproduce an important, well-known statistical relationship of earthquakes.

Malamud et al (1998) analyzed four datasets on the sizes of forest fires and wildfires in the United States and Australia, each of them consisting of 120 to 4284 fires. As illustrated in Fig. 10.4, they indeed found power-law distributions over several orders of magnitude. The estimated scaling exponents of the four distributions fall into a rather narrow range $\tau \in [1.31, 1.49]$.

Power-law statistics or at least heavy-tailed distributions have been recognized in several other wildfire data sets under a broad range of conditions, too (Minnich and Chou 1997, Ricotta et al 1999, Song et al 2001, Malamud et al 2005, Krenn and Hergarten 2009). The majority of the scaling exponents is in the range $\tau \in [1.1, 2.0]$. so that the variablity in the exponents has considerably increased since the study of Malamud et al (1998). In this paper, it was attempted to reproduce the observed variability in τ by a variation of the growth rate θ and the system size. However, all values of τ obtained from this were lower than the theoretical limit value $\tau = 1.19$ discussed in the previous section, so that the result stayed outside the range $\tau \in [1.31, 1.49]$

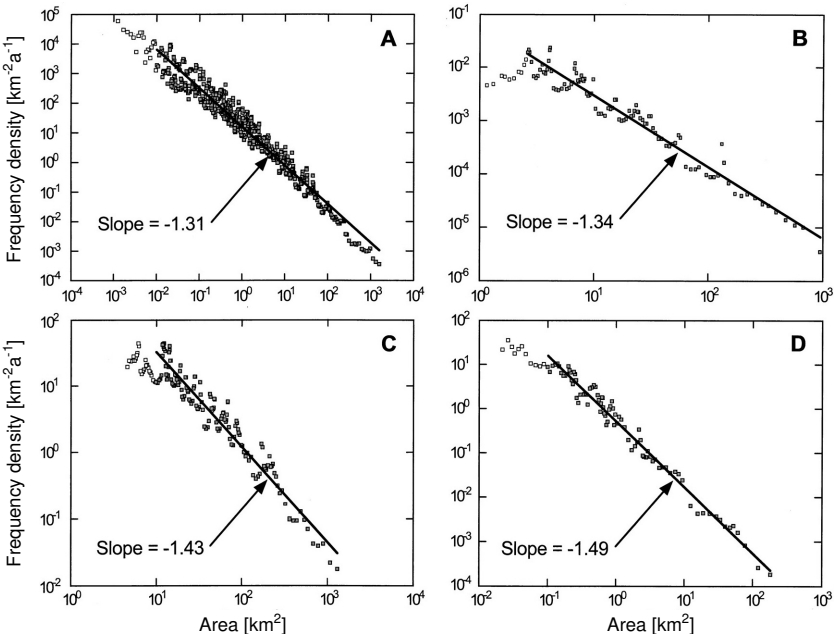


Fig. 10.4: Frequency-area distributions for actual forest fires and wildfires in the United States and Australia (Malamud et al 1998): (A) 4284 fires on U.S. Fish and Wildlife Service lands (1986–1995), (B) 120 fires in the western United States (1150–1960), (C) 164 fires in Alaskan boreal forests (1990–1991), and (D) 298 fires in Australia (1926–1991).

found in this study, and only the lower edge of the wider range $\tau \in [1.1, 2.0]$ can be explained this way.

There are, of course, several reasons to assume that the scaling exponent of the wildfire size distribution indeed varies in nature. It may depend, e.g., on topographic or climatic constraints and particularly on human impact. Fire mitigation may perhaps be the strongest impact, but timber harvesting may be important as well. The question how suppressing small and mid-sizes affects the frequency of large fires seems to be the most controversially discussed topic in this context.

However, it is in general difficult to attribute differences between datasets to a single reason. As an example, Minnich and Chou (1997) and Minnich (2001) found significant differences in the wildfire size distributions of southern California and the adjacent region of Baja California Norte, Mexico. Large fires occur more frequently in southern California, manifested in a scaling exponent $\tau = 1.44$ vs. $\tau = 2.02$ for Baja California Norte. The difference was attributed to the suppression of small fires in California leading to a temporal shift of the fire season towards hot and dry conditions, but the authors

admitted that factors such as flora conservation programs, livestock grazing practices, and wildlife populations may also contribute significantly.

There is at least one data set where the direct human impact by intentional or careless ignition can be delineated. The Canadian Large Fire Database (LFDB) (Canadian Forest Service 2002) contains information on 11,231 fires of more than 2 km^2 in area for the 1959 to 1999 period and distinguishes between fires caused by lightning and human action. 72% of the fires recorded in the LFDB were initiated by lightning, and 25% by humans (the rest is unknown).

Figure 10.5 shows the probability densities of the fire sizes estimated from the LFDB data. Assuming that the deviation of the graphs at large event sizes is due to a cutoff, Krenn and Hergarten (2009) fitted a power-law distribution with $\tau = 1.30$ for all fires (green line). For the lightning-induced fires, $\tau = 1.20$ (blue line) was found which is very close to the theoretical value $\tau = 1.19$ of the forest-fire model. In return, very large anthropogenic fires occur less frequently, manifested in a larger scaling exponent $\tau = 1.61$ (red line). In these data, the human impact is obvious, and an extended forest-fire model addressing the difference between natural and human-made fires will be discussed in the next section.

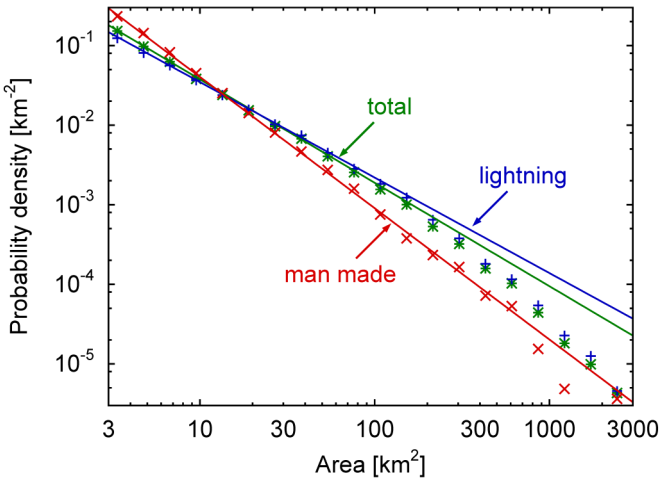


Fig. 10.5: Probability densities derived from the Canadian Large Fire Database (Krenn and Hergarten 2009). Points represent the logarithmically binned data (6.6 bins per decade), straight lines correspond to power laws fitted by a least-squares procedure.

So the forest-fire model predicts the power-law distribution of wildfires at least qualitatively well, and perhaps even the case of natural fires in an undisturbed environment (or whatever the situation with the lowest scaling

exponents may be) quantitatively. But as stated by Zinck and Grimm (2008), it has been largely ignored by fire ecologists. Admittedly, the occurrence of a power-law distribution is one single property, and the scaling exponent is just one number. As, e.g., recently pointed out by Stumpf and Porter (2012), this is in general a rather weak proof of a model's relevance to nature. Apart from this, the forest-fire model appears to be oversimplified. The propagation of real wildfires is, of course, much more complicated than assumed in the model, and it is clearly influenced by temperature, wind, rainfall, and the vegetation itself for each individual fire. However, the forest-fire model was never designed as a model of wildfire dynamics, and its applicability in a short-term regional prediction is clearly limited. Instead, the model rather addresses the susceptibility to large fires at intermediate to long time scales.

Apart from the huge number of parameters disregarded in the forest-fire model, the type of its randomness is often criticized. A random ignition is indeed reasonable in a simplified world. But as pointed out by Zinck and Grimm (2008), the idea that trees pop up randomly, while the propagation of fire is completely deterministic, reverses the typical ecologists view that the vegetation succession is deterministic and the individual fires stochastic. But this is, in principle only a good example of pre-judgment without taking a closer look. In fact, spreading of fire is not really random in nature. If the conditions are good, fires propagate very well, while nothing happens, e.g., under wet conditions. The forest itself contributes to propagation by providing the amount of fuel, and once a fire is ignited, the rest is more or less deterministic. And what about the trees randomly popping on a lattice? As discussed in the previous section, areas are cleared by large fires and then regrow through time in the model. The potential for a new large fire hinges on the mean density of the trees in this area, and it rapidly increases when the density approaches the percolation threshold. So if we do not look at each individual tree (which indeed pops up unrealistically), but consider aggregations of several sites, the model continuously develops fuel to allow large fires. In this sense, the behavior of an individual model tree is as unimportant for large fires as the behavior of each single raindrop for a flood. Motivated by their results on the shape of wildfires (Zinck and Grimm 2008), Zinck et al (2010) designed a coarse-grained version of the forest-fire model where each site corresponds to a larger unit, e.g., a forest stand of several hectares. In their model, the sites do not pop up randomly, but are characterized by a time-dependent susceptibility to burning. This susceptibility is zero immediately after a fire and then increases until a fire destroys this part of the forest. As it can already be expected from the behavior of patches in the original forest-fire model, the results of this model are essentially the same.

Nevertheless, it seems that the model proposed by Ratz (1995) where the role of randomness is just opposite to the forest-fire model is more accepted in forest ecology. This model is also a simple cellular automaton ignoring topography, weather, and details of the forest structure, but fire spread is stochastic, and regrowth or, more precisely, aging, is deterministic. This model does

not reproduce the observed power-law size distribution of wildfires, which was admittedly not in discussion when the model was developed. Instead, the author validated his model using the structural properties of 68 wildfires analyzed by Eberhart and Woodard (1987). These properties mainly refer to the scaling of the outer perimeter of the fires, similarly to Eq. 10.4, and the number and sizes of unburnt islands within burnt areas. In return, Caldarelli et al (2001) concluded from an analysis of three large fires that the Drossel-Schwabl forest-fire model cannot reproduce realistic fires shapes. In contrast, Zinck and Grimm (2008) found that the forest-fire model reproduces the geometric properties of the 68 fires mentioned above after spatial rescaling of the Drossel-Schwabl model as well as the model of Ratz. The result that the accessible perimeter of fires in the forest-fire model is very similar to that of percolation clusters (Eq. 10.4) may lead to the conclusion that these geometric properties of fires are not a very sensitive criterion, so that several models will finally yield good results here. From this point of view the comparison between the Drossel-Schwabl forest-fire model and the model of Ratz ends with the power-law distribution on the side of the Drossel-Schwabl model versus the feeling that the model of Ratz is more realistic from its basic assumptions apparently shared in the scientific community of fire ecology.

10.4 Extensions of the Forest-Fire Model

As discussed in the previous section, the scaling exponents of wildfire statistics in nature show a variation that cannot be explained by the statistical uncertainty alone. The forest-fire model apparently predicts the lower edge of the range found in nature well, while the larger exponents frequently found may arise from specific conditions not captured by the simple model. Most ideas on these specific conditions go into the direction of human impact. So there have been several attempts to extend the forest-fire model in order to predict a variation in the scaling exponent.

The first extension of the model documented in the literature concerns the assignment of an immunity to the sites (Clar et al 1994, 1996). This was implemented in the way that fire spreads from one site to another fails at a given probability g, so that a green tree only catches fire with the probability $1 - g^n$ if n of its neighbors are burning. A simple bifurcation was found with respect to the parameter g: The power-law distribution with the original scaling exponent persists for $g < \frac{1}{2}$, while the power law vanishes and the distribution decays exponentially for $g > \frac{1}{2}$.

Similarly, Hergarten (2002) investigated the forest-fire model with modified rules of fire spreading in order to mimic the influence of climate. Propagation towards the eight nearest and second-nearest (diagonal) neighbors was considered as a representation of a drier climate where fires spread more easily. It turned out that the self-organization towards a power law with the

original scaling exponent is strong enough to compensate this modification. The modified forest-fire model approaches a state where the mean tree density is significantly lower than in the original model (about 0.27 instead of 0.41), but the event-size distribution remains almost the same. Similar results were obtained if only propagation towards two (non-opposite) of the four nearest neighbors is considered, corresponding to a highly anisotropic propagation arising from a preferred direction of wind. Figure 10.6 gives an example of the spatial pattern if propagation is allowed only towards the southern and the eastern neighbor. The anisotropy induced by the wind direction is clearly visible in the pattern, and the strong difference towards the critical state of the original forest-fire model (Fig. 10.1) is obvious. But again, the model just self-organizes to a critical state with a different spatial pattern in order to maintain the fire size distribution with the original exponent τ.

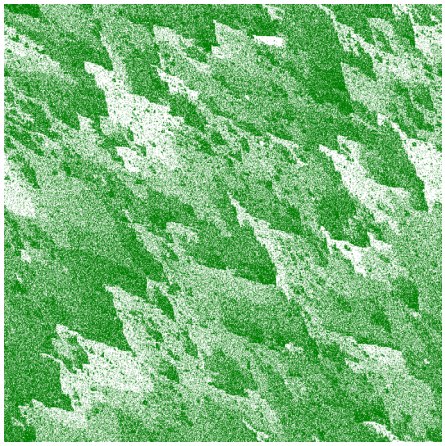


Fig. 10.6: A snapshot from a simulation of the forest-fire model where fires can only spread towards the southern and the eastern neighbor mimicking a preferred wind direction (Hergarten 2002).

In the same study, transient states with respect to the model rules were also considered. This is of particular interest with regard to climate change (e.g., Pueyo 2007, Pueyo et al 2010). The results discussed above suggest that even a much drier climate should not change the power-law distribution of the wildfires. This is, of course, not true during the transition from the critical state with respect to the original rules to the new critical state. This phenomenon was investigated by applying the new rules of fire propagation (i.e., allowing diagonal connections) to the original critical state. So this analysis refers to the fire size distribution immediately after an instantaneous transition to a drier climate. The results shown in Fig. 10.7 suggest that a part of the distribution still follows a power law with the original scaling

exponent, but a heavy tail in the distribution evolves which is clearly outside the power-law. For the parameters considered there ($\theta = 2048$, $L = 8192$) it was found that the mean size of the fires increases by more than a factor 1000 compared to the original model due to the heavy tail of the transient distribution. However, it seems that this research has not been continued systematically since then, although perhaps of particular interest with respect to scenarios of climatic change.

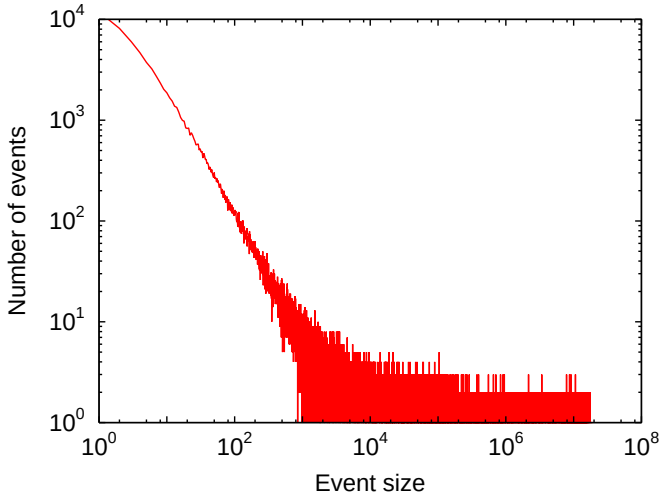


Fig. 10.7: Size statistics of the fires immediately after diagonal connections have been allowed for $\theta = 2048$ on a 8192×8192 lattice (Hergarten 2002).

The differences in the statistics between lightning-induced and anthropogenic fires discussed in the previous section were addressed by the accessible perimeter ignition forest fire model (Krenn and Hergarten 2009). This model hinges on the simple idea that all model trees have the same susceptibility of being ignited by lightning (as it is manifested in the rules of the original forest-fire model), but ignition by human action should be more likely at the border of a forest or at locations where the forest is not very dense. This idea was implemented by assuming that a cluster of trees can be ignited by human action only if a spark hits a site belonging to the accessible perimeter of the cluster (see Fig. 10.3). Allowing ignition of all perimeter sites (the total perimeter) would be even simpler, but as discussed in the previous section, the finite tree density within the cluster would allow ignition almost everywhere then, so that the effect of the restriction of ignition would be almost negligible.

As illustrated in Fig. 10.8, this modification indeed results in a power-law distribution with a larger scaling exponent than the original forest-fire model. While the authors found $\tau = 1.22$ for the original forest-fire model,

the modified model predicts $\tau = 1.51$ for ignition via the accessible perimeter. These results are in very good agreement with the distributions of natural and anthropogenic fires in Canada discussed in the previous section. The data power-law fits in Fig. 10.5 yielded $\tau = 1.20$ for the lightning-induced fires and $\tau = 1.61$ for the anthropogenic fires, so that the agreement with the model data is almost perfect if we take the uncertainty arising from the strong cutoff effects at large fire sizes into account.

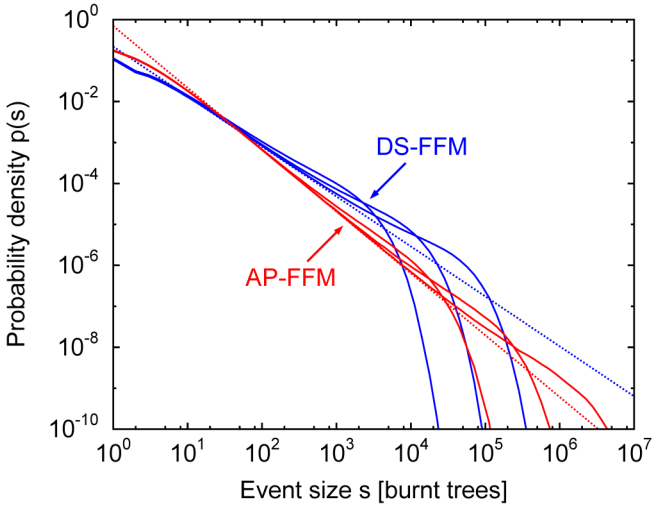


Fig. 10.8: Event-size distributions for the original forest-fire model (DS-FFM, blue) and the accessible perimeter ignition forest-fire model (AP-FFM, red). Curves correspond to growth rates $\theta = 256, 1024$ and 4096 from left to right. The dotted lines indicate the slope for intermediate event sizes estimated by extrapolation to $\theta \to \infty$ (Krenn and Hergarten 2009).

An explanation for the difference in the scaling exponents between the different modes of ignition was given in this study, too. It was found that the size distribution of the existing clusters of trees is almost the same as in the original model, so that the difference directly arises from the difference in the probability of ignition. While the probability that a given cluster of size s is ignited is linearly proportional to s, it scales only like $s^{0.69}$ if ignition is restricted to the accessible perimeter according to Eq. 10.4. This directly leads to a difference of 0.31 in the scaling exponents of the fire size distributions. Even the theory of Hergarten and Krenn (2011) reviewed in Sect. 10.2 can be applied to the accessible-perimeter ignition and leads to $\tau = 1.50$ instead of $\tau = 1.19$ for the original forest-fire model. The result that the size distribution of the existing clusters persists is remarkable because the modifications

discussed earlier introduced changes in existing patterns in such a way that the size distribution of the fires persist.

Beyond this, Krenn and Hergarten (2009) also performed mixed simulations involving both types of ignition. While the occurrence of an overall power-law distribution with a scaling exponent depending on the mixture of the fire types is not surprising, more interesting results were obtained concerning the largest events. As shown in Fig. 10.8, the higher scaling exponent of the anthropogenic fires in absence of natural fires is accompanied by a shift of the large-size cutoff, so that the largest human-induced fires are finally larger than the largest lightning-induced fires. It was found that this result does not hold in a mixed environment. As long as there are enough lightning-induced fires, the long tail of the anthropogenic fires' distribution is cut off, so that the largest fires are finally dominated by lightning. So the occurrence of natural fires prevents the system from developing very large human-made fires which would be in principle even larger than the largest natural fires.

Fire suppression is probably the most controversially discussed human impact. One of the striking questions in this context is whether extinguishing small (and perhaps mid-sized) fires in return may even increase the number of large fires. The simplest idea to mimic fire suppression would be randomly extinguishing burning sites. However, this would be essentially the same as the immunity introduced by Clar et al (1994). This means that moderate fire suppression would not affect the scaling exponent of the size distribution. In return, the large fires would almost vanish if enough effort was spent for fire suppression.

Yoder et al (2011) extended the forest-fire model by a fire resistance depending on the current number of burning sites. Assuming that the resistance decreases with increasing number of burning sites, the fact that it is more difficult to extinguish a large fire than a small fire can be regarded. As a simple realization, the authors assumed that a fire is completely extinguished with a probability

$$P_q(k'_f) = \frac{\epsilon}{k'_f} \tag{10.9}$$

where k'_f is the number of sites currently burning and ϵ is a tuning parameter. In contrast to the original forest-fire model, this rule of extinction does not allow burning down a cluster in one step. Fire spreading is the same as in the first version of the forest fire model where all neighbors of the currently burning sites catch fire in each step. After each step of spreading, it is decided according to Eq. 10.9 whether the fire continues or stops at this stage. If it continues, all neighbors of the burning sites catch fire, and so on. All burning sites remain in this state until the fire has stopped, so that k'_f in Eq. 10.9 is not the size of the fire front at the moment, but the size of the area affected by this fire so far. As Eq. 10.9 is evaluated after the first step of fire propagation where k'_f cannot be larger than 5 even in a dense cluster, ϵ must be lower

than five. Otherwise, all fires would be immediately extinguished after the first step of propagation.

Figure 10.9 shows the results obtained by the authors for $\theta = 4096$ on a 1024×1024 lattice. They obtained power-law distributions with scaling exponents τ depending on the tuning parameter ϵ. For the largest value $\epsilon = 4.7$ considered in this study, they obtained $\tau = 1.74$. The authors observed an almost linear increase of τ with ϵ up to $\epsilon \approx 3.5$, and that τ does not increase further for larger values of ϵ. So their results suggest that variation in fire susceptibility (including fire suppression) may introduce a variation in the scaling exponent from about 1.2 to about 1.74, the largest value found in their study. This range is wider than the difference between natural and anthropogenic fires predicted by the ignition mechanisms discussed above and covers a large part of the variability found in nature (see Sect. 10.3).

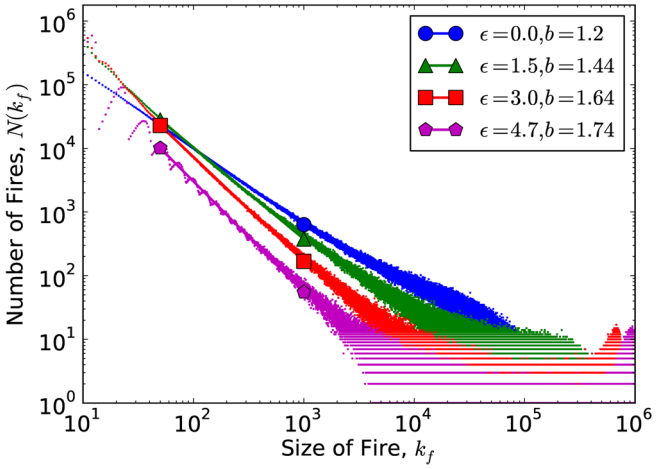


Fig. 10.9: Frequency-size distributions of the model with fire resistance for different values of the tuning parameter ϵ (Yoder et al 2011). The scaling exponent is denoted b in this diagram.

The increase of the scaling exponent with increasing fire resistance suggests that fire suppression would indeed reduce the number of large fires. However, the tails of the distributions in Fig. 10.9 show that the number of large fires in fact increases. With regard to the system size $L = 1024$ used by the authors even a considerable number of system-wide fires occurs for $\epsilon \geq 3$. Similarly to the transition effects discussed earlier in this section, the total loss of trees is dominated by extremely large fires outside the range of the power-law distribution. This would in return imply that fire suppression leads to an extremely increased hazard of large fires. This topic came into discussion even before the idea of SOC was developed (Minnich 1983), but the occurrence

and the significance of this phenomenon in nature seems to be rather complex and might indeed not be captured by such simple models.

In both the accessible perimeter ignition forest-fire model of Krenn and Hergarten (2009) discussed before and the model of Yoder et al (2011) increases in the scaling exponent of the power-law distribution go along with a heavier tail of the distribution. This is not a property of these two particular extensions of the model, but an inherent weakness of the forest-fire model itself. The forest-fire model hinges on the idea that all new trees will be destroyed by fires, otherwise there could be no quasi-steady state. As discussed when deriving Eq. 10.1, $(1-\rho)\theta$ trees must be burnt per ignition event in the mean. If θ is given, this number can only be reduced significantly if the density ρ approaches one. This is exactly how the introduction of immunity by Clar et al (1994) avoids the occurrence of large fires if immunity is large enough. Thus, the forest-fire model with an equilibrium of growth and burning is in principle unsuitable to predict the effects of changes in fire suppression or in fire propagation on the occurrence of large fires.

As mentioned above, the long tail of the distribution of the anthropogenic fires in the accessible perimeter ignition forest-fire model vanishes in a mixed environment of natural and human-induced fires. In principle, two competing mechanisms of burning are present in this case, so that none of them has to maintain an equilibrium with growth alone. In many regions on earth, timber harvesting obviously contributes to the consumption of trees as well as fires. This contribution may be even much higher than the loss of forest by fires. In regions of low ignition, even the loss by naturally dying trees may be significant. Therefore, the basic assumption that burning alone balances the growth of new trees is in principle unrealistic in many environments and may introduce serious artefacts.

A first study on the removal of fuel in the forest-fire model, e.g., by timber harvesting, was recently conducted (Krenn and Hergarten 2013). As a simple idea, the forest-fire model was extended by the rule that ω sites are randomly (or according to a given spatial pattern) selected in each step, and that all trees present on these sites are removed. This simple rule is just the opposite of growing new trees with a parameter ω instead of θ.

Figure 10.10 shows the resulting fire size distributions obtained from this model with random fuel removal. Instead of ω, the parameter $h = \frac{\omega}{\theta}$ which quantifies clearing by harvesting or other processes in relation to growth was used. The scaling exponent of the fire size distribution roughly persists as long as removal is not too strong. In return, the largest fires become smaller since growth is now balanced by burning and removal in sum. Similarly to the model with immunity of Clar et al (1994), the power-law distribution is lost if removal becomes dominant compared to burning. In this regime, the patchy spatial structure occurring in the original forest-fire model is destroyed, and the model behaves like percolation below the percolation threshold. These results may be important for the question in which environments a power-law distribution of forest fires may be expected at all. Apart from this, it opens

the door towards a better understanding of the effect of fire suppression or climatic change. However, a combination with the other extensions of the forest-fire models discussed in this section has not been performed so far.

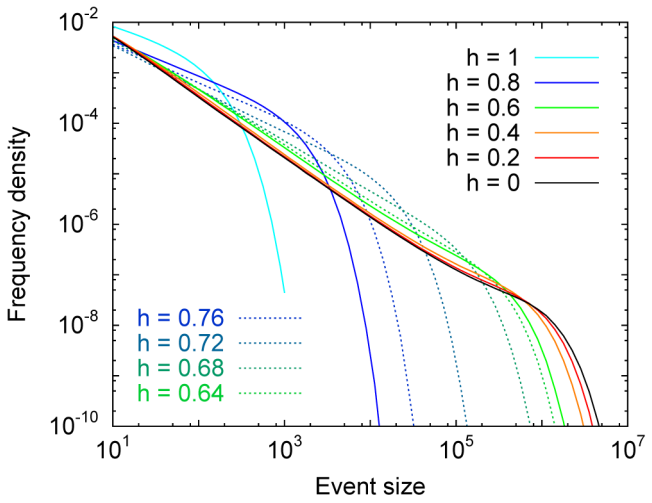


Fig. 10.10: Size distribution of the fires in the model with fuel removal for $\theta = 65,536$ and different values of the clearing parameter h (Krenn and Hergarten 2013).

In sum, the simple forest-fire model is still a fascinating model in the context of SOC after more than 20 years. Despite some theoretical progress concerning the scaling exponent of the size distribution, the understanding of the model's behavior and its relationship to percolation is still incomplete. The forest-fire model is clearly not a model of wildfire dynamics in detail, but rather predicts the susceptibility to large fires at intermediate to long time scales. The interest in this model in the field of fire ecology has obviously grown in the last years, although the discussion is still controversial with good arguments on both sides. But in view of its extensibility in various directions, the forest-fire model should deserve even more attention in the field of fire ecology than it presently receives.

References

Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality. An explanation of 1/f noise. Phys Rev Lett 59:381–384, DOI 10.1103/PhysRevLett.59.381

Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. Phys Rev A 38:364–374, DOI 10.1103/PhysRevA.38.364

Bak P, Chen K, Tang C (1990) A forest-fire model and some thoughts on turbulence. Phys Lett A 147:297–300

Caldarelli G, Frondoni R, Gabrielli A, Montuori M, Retzlaffi R, Ricotta C (2001) Percolation in real wildfires. Europhys Lett 56:510–516

Canadian Forest Service (2002) Canadian Large Fire Database (LFDB), `http://fire.cfs.nrcan.gc.ca/research/climate_change/lfdb/lfdb_download_e.htm`

Christensen K, Flyvbjerg H, Olami Z (1993) Self-organized critical forest-fire model: mean-field theory and simulation results in 1 to 6 dimensions. Phys Rev Lett 71(17):2737–2740, DOI 10.1103/PhysRevLett.71.2737

Clar S, Drossel B, Schwabl F (1994) Scaling laws and simulation results for the self-organized critical forest-fire model. Phys Rev E 50:1009–1018

Clar S, Drossel B, Schwabl F (1996) Forest fires and other examples of self-organized criticality. J Phys: Condens Matter 8:6803–6824, DOI doi:10.1088/0953-8984/8/37/004

Drossel B, Schwabl F (1992) Self-organized critical forest-fire model. Phys Rev Lett 69:1629–1632, DOI 10.1103/PhysRevLett.69.1629

Eberhart K, Woodard P (1987) Distrubution of residual vegetation associated with large fires in Alberta. Can J Forest Res 17:1207–1212

Frette V, Christensen K, Malthe-Sørenssen A, Feder J, Jøssang T, Meakin P (1996) Avalanche dynamics in a pile of rice. Nature 379:49–52

Gabrielov A, Newman WI, Turcotte DL (1999) Exactly soluble hierarchical clustering model: Inverse cascades, self-similarity, and scaling. Phys Rev E 60(5):5293–5300, DOI 10.1103/PhysRevE.60.5293

Grassberger P (1993) On a self-organized critical forest fire model. J Phys A 26:2081–2089

Grassberger P (2002) Critical behaviour of the Drossel-Schwabl forest fire model. New J Phys 4(17):17.1–17.15, DOI 10.1088/1367-2630/4/1/317

Grassberger P, Kantz H (1991) On a forest fire model with supposed self-organized criticality. J Stat Phys 63:685–700

Grossman T, Aharony A (1986) Structure and perimeters of percolation clusters. J Phys A: Math Gen 19(12):L745–L751, DOI 10.1088/0305-4470/19/12/009

Henley CL (1989) Self-organized percolation: a simpler model. Bull Am Phys Soc 34:838

Henley CL (1993) Statics of a "self-organized" percolation model. Phys Rev Lett 71:2741–2744

Hergarten S (2002) Self-Organized Criticality in Earth Systems. Springer, Berlin, Heidelberg, New York

Hergarten S, Krenn R (2011) A semi-phenomenological approach to explain the event-size distribution of the Drossel-Schwabl forest-fire model. Nonlin Processes Geophys 18:381–388, DOI 10.5194/npg-18-381-2011

Honecker A, Peschel I (1997) Length scales and power laws in the two-dimensional forest-fire model. Physica A 239:509–530

Krenn R, Hergarten S (2009) Cellular automaton modelling of lightning-induced and man made forest fires. Natural Hazards and Earth System Sciences 9:1743–48

Krenn R, Hergarten S (2013) The influence of fuel removal in the Drossel-Schwabl forest-fire model. in preparation

Malamud BD, Morein G, Turcotte DL (1998) Forest fires: an example of self-organized critical behavior. Science 281:1840–1842

Malamud BD, Millington JDA, Perry GLW (2005) Characterizing wildfire regimes in the United States. P Natl Acad Sci USA 102(13):4694–4699, DOI 10.1073/pnas.0500880102

Minnich RA (1983) Fire mosaics in Southern California and Northern Baja California. Science 219:1287–94

Minnich RA (2001) An integrated model of two fire regimes. Conservation Biology 15:1549–1553, DOI 10.1046/j.1523-1739.2001.01067.x
Minnich RA, Chou YH (1997) Wildland fire patch dynamics in the chaparral of Southern California and Northern Baja California. Int J Wildland Fire 7(3):221–248, DOI 10.1071/WF9970221
Moßner W, Drossel B, Schwabl F (1992) Computer simulations of the forest-fire model. Physica A 190:205–217
Newman MEJ, Ziff RM (2000) Efficient Monte Carlo algorithm and high-precision results for percolation. Phys Rev Lett 85:4104–4107, DOI 10.1103/PhysRevLett.85.4104
Olami Z, Feder HJS, Christensen K (1992) Self-organized criticality in a continuous, nonconservative cellular automation modeling earthquakes. Phys Rev Lett 68:1244–1247, DOI 10.1103/PhysRevLett.68.1244
Pastor-Satorras R, Vespignani A (2000) Corrections to scaling in the forest-fire model. Phys Rev E 61:4854–4859
Pruessner G, Jensen HJ (2002) Broken scaling in the forest-fire model. Phys Rev E 65:056,707
Pruessner G, Jensen HJ (2004) Efficient algorithm for the forest fire model. Phys Rev E 70:066,707
Pueyo S (2007) Self-organised criticality and the response of wildland fires to climate change. Climatic Change 82:131–161
Pueyo S, Gracia PMLA, Barbosa RI, Cots R, Cardona E, Fearnside PM (2010) Testing for criticality in ecosystem dynamics: the case of Amazonian rainforest and savanna fire. Ecology Letters 13:793–802
Ratz A (1995) Long-term spatial patterns created by fire: a model oriented towards boreal forests. Int J Wildland Fire 5:25–34, DOI 10.1071/WF9950025
Ricotta C, Avena G, M, Marchetti (1999) The flaming sandpile: self-organized criticality and wildfires. Ecol Model 119(1):73–77, DOI 10.1016/S0304-3800(99)00057-5
Schenk K, Drossel B, Schwabl F (2002) Self-organized critical forest-fire model on large scales. Phys Rev E 65:026,135
Song W, Weicheng F, Binghong W, Jianjun Z (2001) Self-organized criticality of forest fire in China. Ecol Model 145(1):61–68, DOI 10.1016/S0304-3800(01)00383-0
Stauffer D, Aharony A (1994) Introduction to Percolation Theory, 2nd edn. Taylor and Francis, London
Stumpf MPH, Porter MA (2012) Critical truths about power laws. Science 335:665–666
Vespignani A, Zapperi S (1998) How self-organized criticality works: A unified mean-field picture. Phys Rev E 57:6345–6362, DOI 10.1103/PhysRevE.57.6345
Yoder M, Turcotte D, Rundle J (2011) Forest-fire model with natural fire resistance. Phys Rev E 83(4):046,118, DOI 10.1103/PhysRevE.83.046118
Zinck RD, Grimm V (2008) More realistic than anticipated: A classical forest-fire model from statistical physics captures real fire shapes. Open Ecol J 1:8–13, DOI 10.2174/1874213000801010008
Zinck RD, Johst K, Grimm V (2010) Wildfire, landscape diversity and the Drossel-Schwabl model. Ecological Modelling 221(1):98–105, DOI 10.1016/j.ecolmodel.2008.12.026

Chapter 11
SOC in Landslides

Stefan Hergarten

Scale invariance at the earth's surface has attracted scientists for several decades. Even the first formal description of scale invariance in nature concerned geomorphic data. In his seminal work, Mandelbrot (1967) measured the length of coastlines with rulers of different lengths and found a power-law relation between the measured length and the ruler's length. The scaling exponent of this power-law relation was used to assign a non-integer dimension between one and two to these lines. The fractional dimension finally led to the term fractals.

In the 1980s, when fractals became popular and subject of several monographs (e.g., Mandelbrot 1982, Feder 1988), geomorphology again provided one of the most fascinating examples: artificial self-affine landscapes (e.g., Voss 1985, Feder 1988). Figure 11.1 shows an example with a local fractal dimension (e.g., Mandelbrot 1985) $D_l = 2.1$, generated by Fourier methods (e.g., Hergarten 2002).

Although these artificial surfaces may be beautiful and even somewhat similar to the real topography of the earth at first, their value turned out to be limited. First, the earth's surface is not perfectly self-similar or self-affine (Evans and McClean 1995) as it is shaped by a variety of processes and shows strong correlations between elevation and slope (e.g., Kühni and Pfiffner 2001) which are not reproduced by simple self-affine surfaces. Furthermore, these surfaces lack important geomorphic elements such as river valleys. And finally, the algorithms behind these surfaces seem to be far away from the present understanding of the tectonic and geomorphic processes shaping the real topography.

A few years after the concept of SOC was introduced, the first attempts to recognize SOC in landform evolution were made (Kramer and Marder 1992, Takayasu and Inaoka 1992, Rinaldo et al 1993). These studies addressed

Stefan Hergarten
Institut für Geo- und Umweltnaturwissenschaften, Albert-Ludwigs-Universität Freiburg, e-mail: stefan.hergarten@geologie.uni-freiburg.de

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

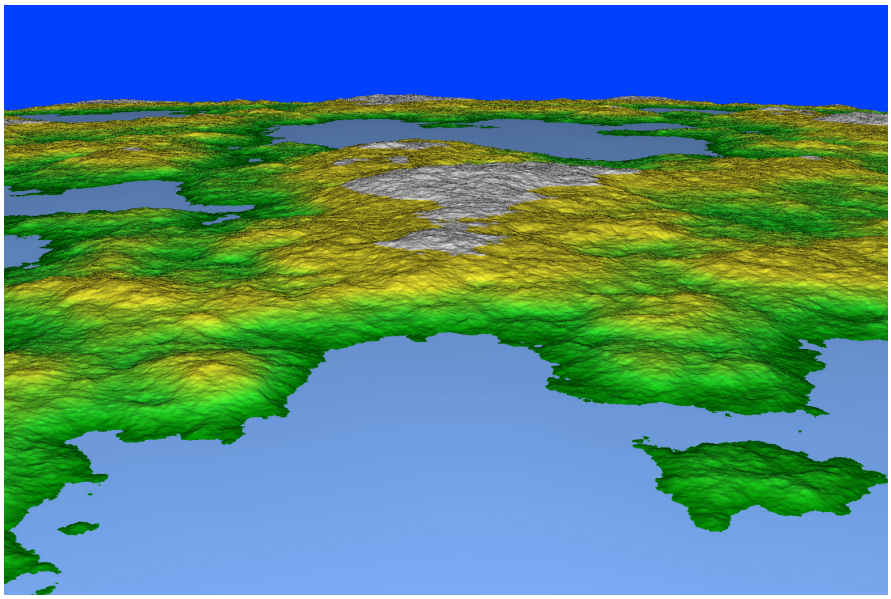


Fig. 11.1: A computer-generated, self-affine surface with a local fractal dimension $D_l = 2.1$. For a more realistic impression, the landscape was flooded up to a certain level and placed on a section of a sphere in order to improve the aerial view.

the statistical properties of river networks using models of fluvial erosion. Scale-invariant properties of river networks were found even before the term fractal was coined (Horton 1945, Strahler 1952, Hack 1957). The models themselves were similar in their spirit. It was assumed that water takes the direction of the steepest descent on the surface, and that the erosion rate of a river segment depends on the discharge and on the local channel slope. The simplest case addresses the evolution of the topography and the river network under constant tectonic uplift where the surface elevation at one or more pre-defined outlet points at the boundary is kept constant.

It was found that the topography including the river network evolves towards a steady state under these conditions. Figure 11.2 shows an example of such a tree-like network. The steady-state networks obtained from these models were found to reproduce several statistical properties of real river networks. Furthermore it turned out that the details of the model, i.e., the constitutive law for the erosion rate as a function of discharge and slope, has a minor effect on the statistical properties of the network.

So these models are examples of self-organization towards a steady state with some scale-invariant properties. Takayasu and Inaoka (1992) even entitled this behavior as a new type of SOC. But as pointed out by Sapozhnikov and Foufoula-Georgiou (1996), this kind of self-organization is not SOC. The

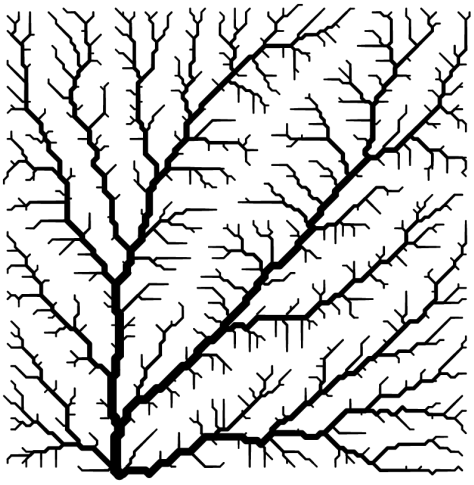


Fig. 11.2: A simulated river network (Hergarten 2002).

evolution ends at a steady state without further fluctuations, so that this state is not critical.

However, tectonic forces and thus the uplift rates are not constant at geological timescales, which is one reason that real landscapes do not achieve a steady state. In order to mimic these permanently changing driving forces, Hergarten and Neugebauer (2001) suggested an extended landform evolution model where the location of the outlet is not constant, but varies through time along the boundary of the model domain. As a consequence, the river network permanently changes and never achieves a steady state. These changes are even reflected by a power law if each change in flow direction at any site is considered as a geomorphic event, and if the event sizes are measured in terms of changes in river discharge or, more precisely, catchment size. But unfortunately, it is impossible to verify this event-size distribution in nature. There is indeed evidence for historical changes in the river network even in mountain belts with deep valleys, but only very few events can be clearly recognized, so that a reasonable statistics seems to be out of reach. Even analyzing the changes in braided river systems which evolve very rapidly compared to large-scale river networks may take more than a human lifetime, and the results obtained from downscaled laboratory experiments (Sapozhnikov and Foufoula-Georgiou 1997) are non-unique with regard to SOC.

So it seems that large-scale landform evolution is in principle unsuitable for recognizing SOC. The rest of this chapter is devoted to landslides which constitute a major natural hazard in almost all mountainous regions and are the presumably the geomorphic process which is most widely studied in the context of SOC.

11.1 Landslide Statistics

Landslides cover an enormous range of scales and a variety of phenomena. In the vast majority of the literature the term landslides is used as a synonym for all gravity-driven mass movements. The involved masses may be either rock fragments or an unconsolidated regolith layer (mainly soil). Depending on the topographic characteristics and the properties of the material, the motion may be dominated by flow, sliding, avalanching, toppling of falling.

The smallest noticeable landslides are rockfalls with a volume in the order of magnitude of 10^{-3} m^3. However, mass movements involving several million cubic meters occur quite frequently. Figure 11.3 shows a rockslide with a volume of about 3×10^7 m^3 that took place in the Matter valley in the Swiss Alps in 1991. Only about 50 years ago, a block of more than one quarter cubic kilometer detached above the Vaiont reservoir in the Dolomite Region of the Italian Alps from a wall and slid into the lake at velocities of up to 30 meters per second. As a result, a wave of water overtopped the dam and swept onto the valley below, with the loss of about 2500 lives. The largest rockslide documented in the European Alps, the Flims rockslide, is even more than 30 times larger with respect to volume than the Vaiont reservoir disaster. Estimates of its total volume cover the range from 8 to 15 km^3 (e.g., von Poschinger 2011).

Fig. 11.3: Debris deposits of a rockslide in the Matter valley (Swiss Alps).

Extensive landslide statistics have been collected for several decades. More than 40 years ago, Fuyii (1969) found a power-law distribution in 650 events induced by heavy rainfall in Japan. In a more comprehensive study, Hovius et al (1997) analyzed about 5000 regolith landslides in the western Southern Alps of New Zealand. Malamud et al (2004) compiled regolith landslide data sets from several regions, each of them consisting of about 1000 to 45,000 events. Some of them were derived from historical inventories, while other consist of events attributed to one triggering event (rapid snowmelt, a rainstorm or an earthquake).

Figure 11.4 shows the frequency density of eight data sets taken from Hovius et al (1997) and Malamud et al (2004) where the area is used as a measure of landslide size. The diagram displays the frequency density, which is simply the product of the probability density and the total number of events. It can be estimated by collecting the objects in (here logarithmic) bins and dividing the number of objects in each bin by the bin width.

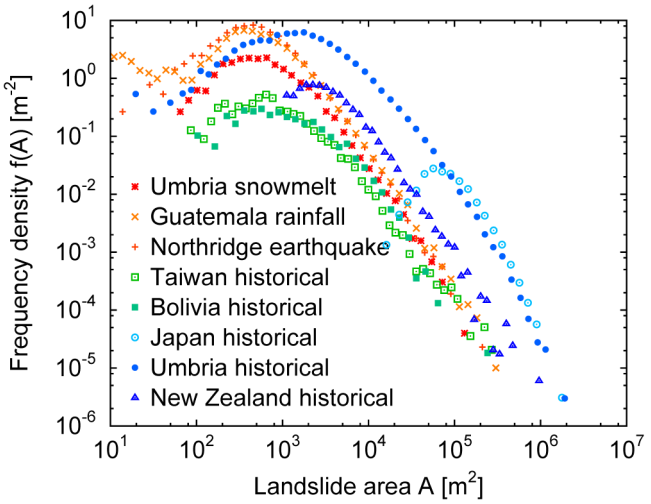


Fig. 11.4: Frequency density of eight landslide data sets (Hovius et al 1997, Malamud et al 2004).

Malamud et al (2004) found a power-law distribution

$$f(A) \propto A^{-\alpha_A} \tag{11.1}$$

with a scaling exponent $\alpha_A \approx 2.4$ at large landslide sizes and a rather small variation between the considered data sets. In particular, there seems to be no systematic difference between the statistics of the landslides triggered by a single earthquake, rainstorm or snowmelt event (red/orange in

Fig. 11.4) and historical inventories involving events arising from various triggers (green/blue).

All datasets displayed in Fig. 11.4 reveal a striking deviation from a power law at small sizes. The rollover of at small sizes indicates a lack of small landslides in all data sets. Although Malamud et al suggested a function to describe it quantitatively, its origin is still unclear, but it seems not to be an artefact of incomplete sampling. This rollover strongly limits the range of landslide sizes where a power law can be found since the largest events are in the order of magnitude of 1 km^2 and thus occur at very low frequencies. As a result, none of the distributions shows a clear power law over more than two decades in area, which is only one decade in linear size. Compared to the distributions of earthquakes (Chapter 9) and wildfires (Chapter 10) this is a rather narrow range.

Available statistics of rock mass movements are much smaller than the inventories of regolith landslides. Malamud et al (2004) reanalyzed three inventories of rockfalls and rockslides originally published by Dussauge et al (2002). Each data set consists of only 89 to 157 events compared to several thousands in the regolith landslide inventories. The frequency densities are displayed in Fig. 11.5. While landslide size was measured in terms of area, volume is used here.

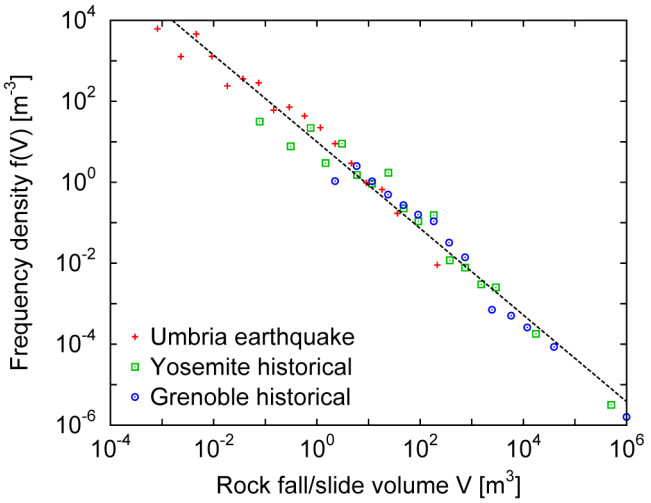


Fig. 11.5: Frequency density of three rockfall inventories (Dussauge et al 2002, Malamud et al 2004). The dashed line illustrates a power law with a scaling exponent of 1.07.

In contrast to the landslide statistics shown in Fig. 11.4, no obvious rollover at small sizes is visible here. Consequently, the power-law distribution extends

over a larger range of scales, about 5 to 7 decades in volume for each data set. Malamud et al found that a power law with a scaling exponent $\alpha_V = 1.07$ fits well to the three datasets without any vertical shift of the curves. This result suggests that the power law even extends over 9 orders of magnitude, but fitting one power-law distribution to a merged data set is in principle dangerous as it strongly depends on the number of events in each data set. However, simultaneously fitting three power-law distributions with the same scaling exponent, but different factors in front of the power law confirms the result $\alpha_V = 1.07$.

As reviewed by Brunetti et al (2009), similar power-law distributions of rockfall and rockslide volumes were found in several other studies. Applying different methods of analysis in different regions on Earth, exponents $\alpha_V = 1.1$ (Guzzetti et al 2003), $\alpha_V = 1.2$ (Guzzetti et al 2004), $\alpha_V \in [1.19, 1.23]$ (Noever 1993), $\alpha_V \in [1.41, 1.52]$ (the original results of Dussauge et al (2002) reanalyzed by Malamud et al (2004), and $\alpha_V \in [1.40, 1.72]$ (Hungr et al 1999) were obtained. Except for two data sets which address rather small scales, all these values fall into the range $\alpha_V \in [1.07, 1.52]$. So the finding that a variation of more than 0.4 in α_V was obtained by applying different methods to the same data sets (Dussauge et al 2002, Malamud et al 2004) suggests that the entire variation in α_V may be a spurious effect of limited statistics.

In addition to an apparent independence on the triggering mechanism, no significant difference between rockfalls and rockslides was revealed. Following the majority of the references cited in this paper, the term rockfalls is therefore used for all types of rapid rock mass movements, in particular rockfalls and rockslides, in the rest of this chapter.

When comparing scaling exponents obtained for rockfalls with those obtained for regolith landslides, we must either transform the area-related regolith landslide distributions to volumes or the volume-related rockfall-size distributions to areas. The simplest assumption is isotropic scaling, $V \propto A^{\frac{3}{2}}$, as used, e.g., by Hovius et al (1997) for regolith landslides. However, anisotropic scaling was revealed in a comprehensive theoretical study by Klar et al (2011), as it was also found much earlier in field studies (e.g., Simonett 1967). Klar et al found a weaker increase of volume with area, $V \propto A^{\gamma}$ with $\gamma \in [1.32, 1.38]$, in very good agreement with field observations. Then, comparing the cumulative distributions with respect to area and volume immediately leads to the relation

$$\alpha_A - 1 = \gamma\,(\alpha_V - 1)\,, \tag{11.2}$$

and thus for $\gamma \leq 1.4$

$$\alpha_V \geq \frac{\alpha_A - 1}{1.4} + 1 \approx 2.0 \tag{11.3}$$

for regolith landslides ($\alpha_A \approx 2.4$). This value is clearly larger than the range $\alpha_V \in [1.07, 1.52]$ found in the rockfall inventories.

In summary, there is growing evidence for power-law size statistics in both regolith landslides and rockfalls (including rockslides). Neither significant regional variations in the scaling exponent nor a dependence on the triggering mechanism has been found. However, the values of the scaling exponent of rockfall size distributions are significantly smaller than those found for regolith landslides.

11.2 Mechanical Models

The power-law distributions found for landslide sizes suggest a relationship to SOC. In the following, the most important modeling approaches in this context are discussed.

All these models address the mobilization of rock or regolith masses. With respect to hazard assessment, this is only half of the story because the runout of a mass movement is as important as the initial mobilization. In particular, debris flows consisting of a mixture of rock and water may travel over distances of several kilometers. Rockfalls, rockslides, and rock avalanches may also differ strongly in their runout behavior. Reviews on models predicting the runout of rock mass movements are given by, e.g., Dorren (2003) and Volkwein et al (2011). In principle the runout may affect the size distribution of the landslides, too. First, rockfall and rockslide inventories mostly refer to the deposited volume which is in general larger than the detached volume as the compactness of the material decreases during the movement. And second, the volume may also increase due to the entrainment of further rock masses. However, dilatancy should not vary strongly with the event size, and the second effect becomes significant for a small class of mass movements only. Under these aspects it seems to be reasonable that the models attempting to relate landslides to SOC only address the detached volumes or the related areas.

The stability of slopes and cliffs is a mechanical problem involving stresses mainly induced by gravity, but in some cases also by variations in temperature or pore water pressure. In rock, pre-defined fracture patterns may be of particular importance, while existing zones of weakness (e.g., clay layers) may strongly affect the stability of a regolith layer. The topography defines the boundary condition for the three-dimensional mechanical problem and is thus at least as important as the mechanics inside the domain. Although this is all clear on a qualitative level, and even the differential equations behind it seem to be well-known, the question whether the power-law distribution arises from an evolving fracture pattern (in rock), another type of stress redistribution or the topography changing through type is still open.

In this section, two models addressing the redistribution of stress are discussed. Both are reduced to two dimensions and in principle concern the stability along a pre-defined slip surface as it is often found in regolith land-

slides. Topography is not directly considered in these models. In their spirit, both models can be seen as extensions of the classical limit equilibrium approach going back to W. Fellenius in the 1920s. In this approach, the stability of a given slip surface is characterized by the factor of safety FS that is defined as the ratio of maximum shear stress $\tau_{\max}$ where the material remains stable and the actual shear stress τ,

$$FS = \frac{\tau_{\max}}{\tau}. \tag{11.4}$$

So the slope remains stable as long as $FS \geq 1$.

The models discussed in the following extend this approach by progressive failure using a local factor of safety. If $FS < 1$ at any location, local failure occurs and leads to an increase of τ and thus to a decrease of FS in the neighborhood. This idea is basically the same as the idea behind the Burridge-Knopoff earthquake model (Burridge and Knopoff 1967) and its most widespread cellular automaton version, the Olami-Feder-Christensen (OFC) model (Olami et al 1992). Although 20 years old, the OFC model is still one of the most widely studied models in the field of SOC, and it is discussed in detail in almost all books on SOC (e.g., Bak 1996, Jensen 1998, Hergarten 2002). Figure 11.6(a) illustrates a physical realization of the OFC model. A set of blocks on a regular lattice is interconnected by springs and held by static friction at the ground. The force $u_{i,j}$ acting on each block increases through time due to an additional connection with a rigid upper plate moving at a constant velocity. When the force acting on a site reaches the limit of friction, the site becomes unstable and is immediately displaced to a new position characterized by zero total force. As a result, a fraction α of the force $u_{i,j}$ is transferred to each of the four nearest neighbors, leading to the relaxation rule of the OFC model

$$u_{i\pm1,j} := u_{i\pm1,j} + \alpha\, u_{i,j}, \quad u_{i,j\pm1} := u_{i,j\pm1} + \alpha\, u_{i,j}, \quad \text{and} \quad u_{i,j} := 0. \tag{11.5}$$

Here, the symbol := means that the value of the variable is replaced with the value at the right-hand side. As a part of the force is transferred to the upper plate (depending on the strength of the springs), α must be smaller than one quarter, making the model nonconservative.

The long-term driving introduced by a rigid upper plate in the earthquake model mimics the long-term displacement between the walls of a geological fault or a subduction zone. Such a way of driving is obviously absent in case of landslides, so that the driver plate has been removed in the realization shown in Fig. 11.6(b). Instead, the blocks connected by elastic springs have been placed on an inclined surface, resulting in a constant driving force in downslope direction.

But apart from the different way of long-term driving, the absence of the rigid driver plate also affects the rule of relaxation in case of local instability. In the earthquake model, the relaxed stress is redistributed among the

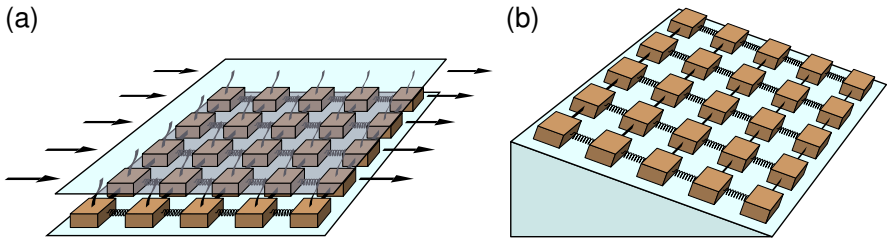


Fig. 11.6: (a) Geometric representation of the Olami-Feder-Christensen earthquake model. (b) Transfer of the idea to simulate progressive slope failure at a given slip surface.

four nearest neighbor sites and the driver plate, resulting in a nonconservative relaxation rule (Eq. 11.5). This property is crucial for reproducing the size-distribution of real earthquakes as the scaling exponent of the event-size distribution roughly approaches 1.2 (e.g., Hergarten 2002) in the conservative limit, which is much too low. Without the driver rigid upper plate, the redistribution of stress becomes conservative, and then the scaling exponent of about 1.2 is clearly too far off from the value $\alpha_A \approx 2.4$ found for regolith landslides.

As the conservative character of the model stems from the principle of conservation of momentum, Hergarten and Neugebauer (2000) looked for a way to obtain larger scaling exponents under conservative stress transfer. They extended the model by a component of time-dependent weakening, which means that the threshold of instability decreases through time between events and is reset after each event.

The model involves two local variables $u_{i,j}$ and $v_{i,j}$ defined on a square lattice. The variable $u_{i,j}$ relaxes conservatively in case of instability like the stresses in the model shown in Fig. 11.6(b) would do. In return, $v_{i,j}$, describing time-dependent weakening, is locally reset to zero in case of instability without any transfer to the neighbors:

$$u_{i\pm1,j} := u_{i\pm1,j} + \frac{1}{4}\,u_{i,j}, \quad u_{i,j\pm1} := u_{i,j\pm1} + \frac{1}{4}\,u_{i,j}, \quad u_{i,j} := 0 \text{ and } v_{i,j} := 0 \,. \tag{11.6}$$

Between events, both variables increase at given rates:

$$\frac{d}{dt}u_{i,j} = r_u \quad \text{and} \quad \frac{d}{dt}v_{i,j} = r_v. \tag{11.7}$$

The rate r_u may describe an increase of stress due to long-term changes in topography, while r_v quantifies the rate of weakening through time. It was assumed that slope stability depends on the product of both variables, i.e., that a site becomes unstable if

$$u_{i,j}\, v_{i,j} \geq 1. \tag{11.8}$$

So the model can be directly transferred to the factor of safety approach (Eq. 11.4) by defining $u_{i,j} = \tau$ and $v_{i,j} = \frac{1}{\tau_{\max}}$ locally. The latter means that, as long as the slope remains stable, the threshold shear stress $\tau_{\max}$ decreases like $\frac{1}{t}$ where t is the time since the last instability at this location. This is, of course, just an ad hoc assumption.

A power-law distribution of the event sizes with a scaling exponent close to 2 was found, and it was theoretically shown that this exponent is independent of the driving rates r_u and r_v. The value $\alpha \approx 2$ was within the range of estimates of α_A for regolith landslides at that time and even in fair agreement with the apparently most reliable value $\alpha_A \approx 2.4$ suggested by Malamud et al (2004).

The model suggested by Piegari et al (2006a,b) is similar in its spirit, but differs in some details and, more importantly, concerning its dissipative character. The local variable is the inverse of the factor of safety (Eq. 11.4) and thus simply proportional to the local shear stress. Reasoned by the existence of several dissipative components in landsliding processes, such as evaporation of water or volume contractions, the authors skipped the conservation of stress and allowed an arbitrary degree of dissipation in the relaxation rule.

At this level, the model is just another physical interpretation of the OFC earthquake model. In extension of the OFC model, the authors replaced the infinitely slow long-term driving with a finite driving rate, as it was investigated by Hamon et al (2002) in the context of solar flares. Furthermore, they introduced an anisotropic relaxation rule since stress transfer in direction of the slope may be stronger than perpendicular to the slope.

Power-law distributions for the event sizes were found for several combinations of the model parameters (dissipation, anisotropy, and driving rate), resulting in a rather large range of scaling exponents. The range includes the values found for regolith landslides in nature. In some cases, the power-law distribution was even lost. In a more recent paper (Piegari et al 2009), the results of this model were quantitatively compared to some of the landslide inventories discussed in the previous section. The authors found combinations of the model parameters that reproduce both the scaling exponent and the rollover at small landslide sizes quite well after spatial scaling of the discrete, nondimensional model. However, the choice of the parameters and the spatial scale is only based on a fit to the data, and there seems to be no way to support this choice from physical principles so far. So it is still unclear why the model parameters should always be in a rather narrow range to yield similar landslide size distributions under strongly different conditions.

Despite the promising results obtained from the two models discussed in this section, some critical comments should be made. The first one mainly concerns the range of scaling exponents obtained from the model of Piegari et al that is obviously much wider than the variation found in nature. This problem also concerns the original OFC model with respect to real earth-

quakes. In principle, the degree of dissipation introduces a tuning parameter which cannot be constrained using physical arguments.

The second criticism arises from the existing knowledge on the behavior of the OFC model. The occurrence of nearly periodic large events was soon discovered (Olami and Christensen 1992), and recently a more or less complete understanding how the power-law distribution arises from the synchronization and the desynchronization of patches toppling almost periodically has been achieved (Hergarten and Krenn 2011). The organization towards an apparently critical state extends over many periods, so that these findings may even support the arguments against the applicability of the OFC model to real earthquakes. With respect to landslides, this argument may even be more severe as it is very difficult to imagine that the power-law distribution emerges after a long series of almost periodic sliding events involving parts of a slope. Due to its similarity with the OFC model, it can be expected that the nonconservative model of Piegari et al behaves exactly like this. For the two-variable model of Hergarten and Neugebauer, there seems to be no further knowledge on its organization towards a critical state, but a similar behavior might be expected.

Finally, the question whether these mechanical models refer to the statistical distributions derived from landslide inventories at all should be taken into account. Since changes in topography are neglected in these models, they describe slip events with a small displacement on an individual slope. Such events have been subject of research in the last decades and may finally help to understand landslide dynamics or even help to predict large landslides, but the events recorded in landslide inventories take place on the landscape scale. Apparently very little is known about the size statistics of these small slip events, but even if they are power-law distributed it is not clear whether this distribution has any relation to the landslide distribution on the landscape scale.

11.3 Geomorphic Models

The second class of models attempting to relate landslides to SOC is part of the large group of landform evolution models. In these models, stresses in the material are not explicitly considered. Instead, slope instability is assumed to depend on properties of the relief, mainly on local slope. In a simplified view, these model approaches can be characterized by the key word sandpile dynamics. Apart form numerical modeling, this topic was also addressed in several laboratory experiments with different granular materials (e.g., Frette et al 1996, Katz and Aharonov 2006, Juanico et al 2008).

In the context of sandpile dynamics, the Bak-Tang-Wiesenfeld (BTW) model (Bak et al 1987, 1988) which was the first model of SOC and still seems to be some kind of paradigm should be mentioned first. This model is

often denoted sandpile model, and even the entire class of models which are similar in their rules are often referred to as sandpile models.

For the two-dimensional BTW model, the presumably most reliable estimate on the scaling exponent of the avalanche size distribution in the limit of infinite system size is $\alpha = 1.27$ (Chessa et al 1999). Interestingly, this value is almost in the middle of the range $\alpha_V \in [1.07, 1.52]$ found for rockfalls and rockslides in nature. And as it should be no problem to accept sandpile dynamics as a simplified representation of rockslides, the problem of relating rockslides (and in principle rockfalls, too) to SOC seems to be solved.

However, the problem in this reasoning is not the question whether sandpile dynamics captures the processes relevant for rockslide dynamics, but the relationship between the BTW model and sandpile avalanches. In the BTW model, a site becomes unstable if its local variable becomes too large. This local variable is often considered as a number of grains at this site and may thus be seen as a representation of surface height. In contrast, the stability of a sandpile depends on the local slope gradient, which should be related to differences in the numbers of grains at neighbored sites instead of their absolute number. Furthermore, redistribution of grains in case of instability should not be isotropic as it is in the BTW model, but mainly in downslope direction.

To get around this fundamental problem, one may be tempted to skip the idea that the variable in the BTW model represents a number of grains, but interpret it as an abstract property that is somehow related to the slope of a sandpile. However, the attempt to relate this variable to slopes succeeds in one dimension, but quantitatively fails on a two-dimensional lattice (Hergarten 2002, 2003). So the BTW model provides a fundamental description of avalanche propagation on a rather abstract level, but a physically consistent relation to sandpile dynamics or any type of gravity-driven mass movements is not visible.

The presumably first geomorphic models to reproduce power-law statistics in landslide dynamics were published 15 years ago (Densmore et al 1998, Hergarten and Neugebauer 1998). Compared to the most widespread models in the field of SOC, these models are rather complicated and involve several parameters.

The model of Hergarten and Neugebauer (1998) is based on partial differential equations. It contains two variables, the surface elevation $H(x_1, x_2, t)$ and the thickness of an upper mobile layer $\kappa(x_1, x_2, t)$, both being functions of the horizontal coordinates x_1 and x_2 and the time t. The material in the mobile layer flows at a velocity proportional to the slope of the surface if a given threshold is exceeded. This behavior is represented by the differential equation

$$\frac{\partial H}{\partial t} = \mathrm{div} \begin{cases} \alpha \left(\kappa |\nabla H| - \beta\right) \frac{\nabla H}{|\nabla H|} & \text{if} \quad \kappa |\nabla H| > \beta \\ 0 & \text{else} \end{cases} . \tag{11.9}$$

The parameter α is related to the flow velocity at given slope, while β defines the threshold where flow starts. The symbols div and ∇ refer to the two-dimensional divergence and gradient operators, respectively. It was further assumed that material from the lower solid layer becomes mobile at a (spatially and temporally) random rate r which is the only random component in the model. In return, the thickness of the mobile layer decays with a given time constant τ. Furthermore, the entrainment of further material due to flow was taken into account. These phenomena were incorporated by the second differential equation

$$\frac{\partial}{\partial t}(H-\kappa) = -r + \frac{\kappa}{\tau} \underbrace{-\gamma\alpha\left(\kappa|\boldsymbol{\nabla} H| - \beta\right)|\boldsymbol{\nabla} H|}_{\text{if } \kappa|\boldsymbol{\nabla} H|>\beta} \tag{11.10}$$

where the additional parameter γ quantifies the entrainment of material by flow. The model was applied to individual slopes, and long-term driving was introduced by a constant rate of lowering at the toe of the slope mimicking the incision of a river.

At that time, the system of differential equations could only be solved with reasonable effort on lattices of no more than 64×64 sites. A power-law distribution of the landslide sizes was found over only one and a half order of magnitude in area. The authors analyzed cumulative distributions. Transferred to non-cumulative frequency densities, they obtained a scaling exponent $\alpha_A \approx 2.1$ which is not far off from the values found for regolith landslides. However, serious parameter studies have not been performed. So the question remains whether this model predicts a universal scaling exponent or whether there is a significant dependence on the model parameters. Apart from this, the geometry of the events seems not to be very realistic. As illustrated in Fig. 11.7, the landslides are rather long and tall and look even a little like gorges in direction of the slope.

The model of Densmore et al (1998) is a rather comprehensive landform evolution model where landsliding is only one component beside fluvial sediment transport and diffusive slope processes. As a major difference towards the models discussed before, slope instability is not treated as a progressive phenomenon. Only the initiation of a landslide at any location is considered, while the size of the resulting event is completely determined by the existing topography and by an ad hoc rule. This topography is, in return, the result of all the processes considered in the model, including previous landslides.

Compared to the model discussed above, the authors attempted to include more knowledge on the stability of real slopes instead of using ad hoc rules for the initiation of landslides. Most of this knowledge hinges on the concept of the factor of safety discussed in the previous section (Eq. 11.4) in combination with the Mohr-Coulomb failure criterion. This criterion is widely used in mechanics and states that failure occurs if the shear stress exceeds the maximum shear stress given by

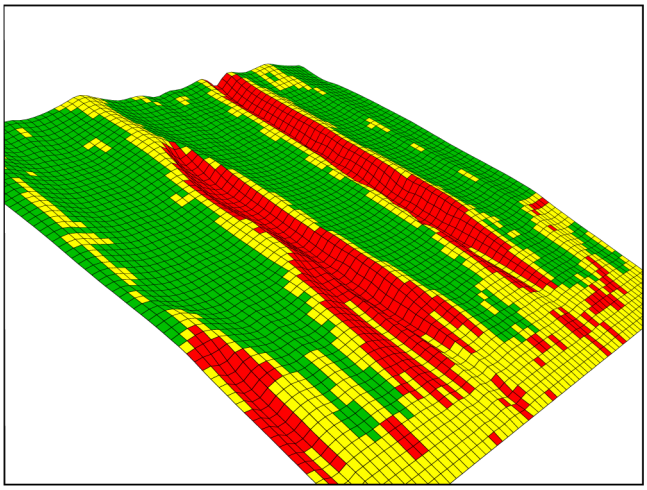


Fig. 11.7: Surface after a large landslide in the model of Hergarten and Neugebauer (1998). Regions that were unstable are yellow; regions where a significant loss of height occurred (larger than the incision of the river) are red.

$$\tau_{\max} = \sigma \tan\phi + C \tag{11.11}$$

where σ is the normal stress. The parameters ϕ and C describe the properties of the material where ϕ is the angle of internal friction and C is the cohesion. For the simplest case of a layer of constant thickness d on a potential failure plane inclined by an angle θ, the Mohr-Coulomb criterion immediately leads to

$$FS = \frac{\tan\phi}{\tan\theta} + \frac{C}{\rho g d \sin\theta} \tag{11.12}$$

where ρ is the density and g is the gravitational acceleration. This simple relationship is often used as a first estimate. It states that planes with angles of inclination $\theta < \phi$ are always stable, while cohesion even enables steeper planes to remain stable as long as the layer is thin.

Densmore et al used this criterion to discriminate sites where landslides may be initiated and the maximum landslide volume at these locations. As they wanted landslides to be initiated only close to the toe of hillslopes, they searched the lowest pair of neighbored sites at each hillslope where the slope angle β between both is larger than ϕ. In Fig. 11.8 these two sites are colored yellow. In the next step, the authors used the Mohr-Coulomb criterion to estimate the maximum height difference H_c between these sites where the slope remains stable and related it to the actual height difference H. They assumed that failure occurs at a probability

$$p = \frac{H}{H_c} + rt \tag{11.13}$$

where t is the time since the last event at this site, and r gives the rate of increase in probability due to time-dependent weakening. In case of instability, a potential landslide volume is computed. For this, they made an estimate of the most likely plane of failure using the Mohr-Coulomb criterion and found that it dips at the angle $\theta = \frac{\beta+\phi}{2}$. The volume above this plane, colored red in Fig. 11.8, defines the maximum possible volume of a landslide at this hillslope. Based on their own empirical results (Densmore et al 1997), they finally assumed that the real landslide volume is directly proportional to the time since the last landslide initiated at this location, limited by the maximum volume.

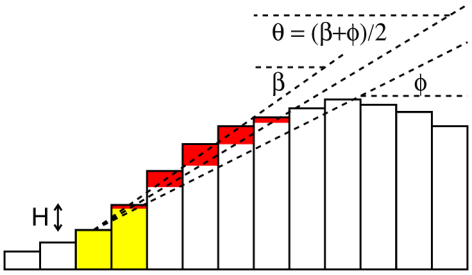


Fig. 11.8: Illustration of the landsliding algorithm suggested by Densmore et al (1998). Landslides can be initiated at the lowest pair of sites on a hillslope where the slope angle β exceeds ϕ (yellow). The maximum landslide volume (red) is defined by a plane that dips at an angle $\theta = \frac{\beta+\phi}{2}$.

Comparing the physical basis of the model with the number of ad hoc rules raises the question whether the physically-based part of the model has any effect on the results. But apart from this, the derivation of the most likely dip angle θ and the maximum stable height difference H_c are wrong. The authors considered the height difference H and the slope angle β between the considered sites as independent and claimed a quite large degree of freedom when deciding which one is variable and which one is given by the actual topography. So it is not surprising that their result on the maximum stable slope is not in agreement with the simple estimate given by Eq. 11.12.

Taking these aspects into account, the part of the model referring to landslides is just a combination of ad hoc rules, similarly to the model of Hergarten and Neugebauer discussed above, but more complicated. However, it should be kept in mind that such rules are not necessarily bad as long as they are reasonable and the results make sense. Similarly to Hergarten and Neugebauer, Densmore et al obtained power-law distributions of the landslide sizes within a narrow range of scales. In two simulations involving different strength of the

material (represented by ϕ and C) they found values $\alpha_V = 2.2$ and $\alpha_V = 1.8$ with respect to the volume over about one order of magnitude. These values are in very good agreement with the estimate $\alpha_V \approx 2.0$ (or slightly larger, depending on the scaling between volume and area) for the volumes of real regolith landslides given in Eq. 11.3. The smaller scaling exponent occurred at higher strength, and this result goes even in the direction that the scaling exponent for rockfalls and rockslides is smaller than that of regolith landslides. However, this may also be a matter of fitting straight lines over rather narrow ranges, and if any error bars had been given, they would surely be larger than the difference between the two values.

After several years of apparent silence in this field, a new approach focusing on rockfalls (again including rockslides) has been recently published by Hergarten (2012). This model is inspired by ideas on sandpile dynamics and extremely simple compared to the other models reviewed in this chapter. In return it is, however, more or less completely based on ad hoc rules. The basic assumption is that landslides can in principle be triggered at any site with a probability that depends on the local slope gradient. All other contributions to rock instability in nature such as fracturing are mimicked by the randomness of the triggering process.

In analogy to the fluvial erosion models mentioned in the introduction, the gradient at each site is computed in the direction of steepest descent among the eight (direct and diagonal) neighbors on a rectangular lattice, what is called D8 algorithm (O'Callaghan and Mark 1984). It is further assumed that slopes below a lower threshold slope $s_{\min}$ remain stable under all conditions, while slopes above an upper threshold slope $s_{\max}$ are destabilized by any impact. For slopes s between $s_{\min}$ and $s_{\max}$ a linear increase of the probability of instability in case of an impact is assumed:

$$p = \frac{s - s_{\min}}{s_{\max} - s_{\min}}. \tag{11.14}$$

If a site becomes unstable, material is removed until its slope decreases to $s_{\min}$. The downslope motion of unstable rock masses and their deposition is not computed, only the volume of detached material is recorded and used for the event size statistics. The effect of the event on its vicinity, i.e., progressive destabilization in the source area of the rockfall, is mimicked by exposing the eight neighbored sites to the same random impact as the unstable site, so that each of them may become unstable with a probability given by Eq. 11.14, too. Those sites which were triggered without becoming unstable are assumed to be stable at their present slope and cannot be destabilized by further impacts unless their slope increases as a consequence of further removal of material at neighbored sites. This is realized by replacing $s_{\min}$ of these sites by the present value of s.

In contrast to all the models discussed earlier, this model only simulates the occurrence and the size of rockfalls on a given relief. Long-term driving forces, mainly fluvial erosion in combination with tectonic uplift and, particularly

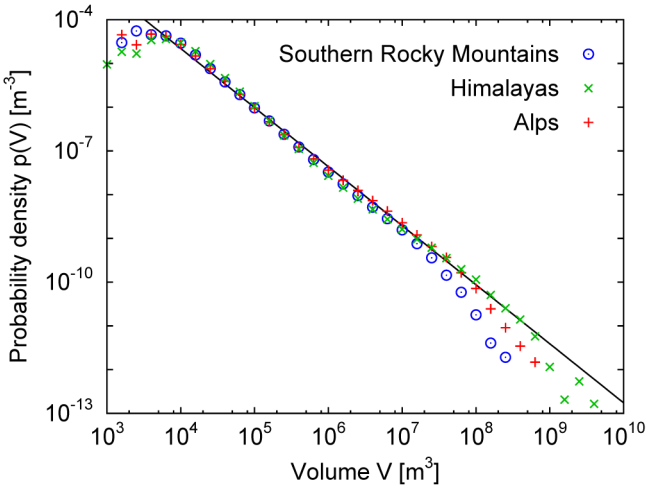


Fig. 11.9: Probability density of the rockfalls predicted by the model of Hergarten (2012) for the European Alps (43–48° N and 5–16° E), the central part of the Himalayas (26–31° N, 82–92° E) and for the southern part of the Rocky Mountains (35–45° N and from 105° W to the West Coast), computed with $s_{\min} = 1$ and $s_{\max} = 5$. The straight line corresponds to a power-law distribution with an exponent $\alpha_V = 1.35$.

important in the context of rockfalls, glacial erosion, are not considered. So if this model yields a power-law distribution of the rockfalls, it only shows that the relief it is applied to has critical properties with respect to this mechanism. In the context of SOC, this is clearly a disadvantage, but in return it might allow a hazard assessment for a given region which is not so easy with models bringing their own mechanism of long-term driving.

The model was applied to Digital Elevation Models of three mountain belts: the European Alps, the central part of the Himalayas, and the southern part of the Rocky Mountains. The elevation data were taken from the ASTER Global Digital Elevation Model (a product of METI and NASA) with a resolution of 1 arc second, corresponding to about 20–30 m.

As illustrated in Fig. 11.9, the model predicts a power-law distribution with a scaling exponent $\alpha_V = 1.35$ for all three regions, although they strongly differ in their topographic characteristics. This value falls perfectly into the range $\alpha_V \in [1.07, 1.52]$ found for rockfalls and rockslides. Significant differences between the regions only concern the cutoff behavior at large event sizes. The results shown in Fig. 11.9 were obtained using the parameter values $s_{\min} = 1$ and $s_{\max} = 5$, only justified by the rule of thumb that the majority of real rockfalls and rockslides occurs at slope angles greater than 45°. However, it was shown that a variation of the model parameters $s_{\min}$ and $s_{\max}$ within a reasonable range has a minor effect on the scaling expo-

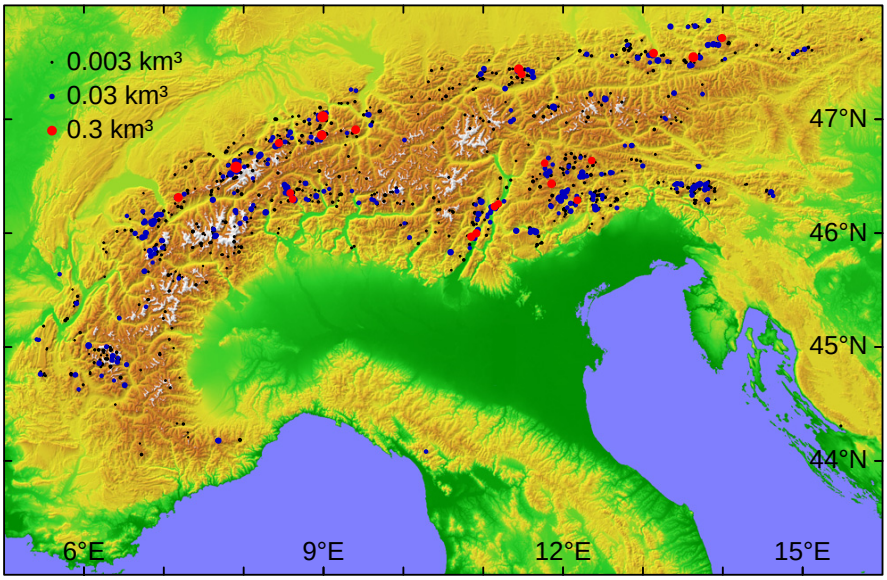


Fig. 11.10: Rockfalls with $V \geq 10^{-3}$ km^3 predicted for a 2000 year time span in the Alps (Hergarten 2012). Black: $V \in [0.001, 0.01)$ km^3 (756 events), blue: $V \in [0.01, 0.1)$ km^3 (301 events), red: $V \geq 0.1$ km^3 (21 events).

nent of the event-size distribution. Similarly to the differences between the considered regions, variations in the parameters mainly affect the cutoff behavior at large event sizes. The regional differences in the cutoff behavior were interpreted in terms of subcriticality of the present relief with respect to the model's mechanism. It was concluded that the Himalayas are closer to a critical state than the Alps, which are themselves closer to a critical state than the southern part of the Rocky Mountains.

In the same paper, a first attempt do derive a topography-based rockfall hazard map from the model was also made. The map presented in Fig. 11.10 is based on a prediction of a 2000 year time span and shows a rather inhomogeneous distribution of the hazard in the European Alps. The largest predicted event is illustrated in Fig. 11.11. It involves a volume of about 0.5 km^3 and is predicted to occur with a rather high probability of one per 500 years. However, it was already admitted in the original paper that quantitative assessments based on this model must be treated with some caution. First, assigning an absolute time scale to the model is rather uncertain. And second, variations in the parameters $s_{\min}$ and $s_{\max}$ which can only be guessed so far have a stronger influence on the largest events than on the power-law distribution itself. It was already discussed in the original paper that even a small increase in $s_{\min}$ and $s_{\max}$ by 20 % reduces both the size and the probability of occurrence of the largest events by a factor of two. As even stronger

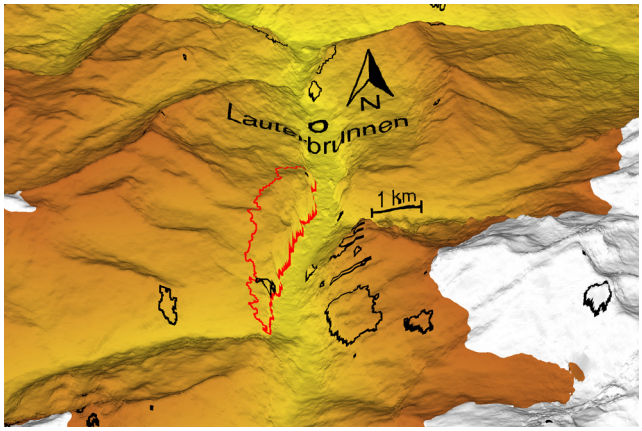


Fig. 11.11: The largest event predicted for the Alps (Lauterbrunnen valley, $V \approx 0.5$ km^3, red). The black lines correspond to smaller events predicted for a 2000 year time span.

regional variations in these parameters can be expected due to lithology, estimating the size and frequency of the largest events in a mountain belt seems to be rather uncertain. So this model may provide a tool for hazard assessment, but any serious application requires additional data that cannot be derived from physical principles in a straightforward way.

Nevertheless, the model seems to have a large potential for both application to rockfall hazard assessment and for clarifying the role of SOC in rockfall dynamics. But as mentioned above, the latter first requires an extension of the model by long-term driving processes such as fluvial or glacial erosion that locally steepen the relief and thus supply the potential for mass movements.

To summarize, there is growing evidence for power-law size distributions in different types of landslides. The scaling exponents found for regolith landslides strongly differ from those found for rockfalls and rockslides, but each of this classes may be characterized by a universal scaling exponent. A handful of models has been designed to reproduce these power-law distributions. Most of them address regolith landslides, an all hinge on ad hoc rules. So far none of them provides a consistent explanation for the difference in the scaling exponents found for different types of landslides. Even none of them can uniquely identify any type of landsliding as a phenomenon governed by SOC, but this applies to almost all natural phenomena considered in the context of SOC.

References

Bak P (1996) How Nature Works – the Science of Self-Organized Criticality. Copernicus, Springer, Berlin, Heidelberg, New York

Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality. An explanation of 1/f noise. Phys Rev Lett 59:381–384, DOI 10.1103/PhysRevLett.59.381

Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. Phys Rev A 38:364–374, DOI 10.1103/PhysRevA.38.364

Brunetti MT, Guzetti F, Rossi M (2009) Probability distribution of landslide volumes. Nonlin Processes Geophys 16:179–188

Burridge R, Knopoff L (1967) Model and theoretical seismicity. Bull Seismol Soc Am 57:341–371

Chessa A, Stanley HE, Vespignani A, Zapperi S (1999) Universality in sandpiles. Phys Rev E 59:R12–R16

Densmore AL, Anderson RS, McAdoo B, Ellis MA (1997) Hillslope evolution by bedrock landslides. Science 275:369–372

Densmore AL, Ellis MA, Anderson RS (1998) Landsliding and the evolution of normal-fault-bounded mountains. J Geophys Res 103:15,203–15,219

Dorren LKA (2003) A review of rockfall mechanics and modelling approaches. Progress in Physical Geography 1:69–87

Dussauge C, Helmstetter A, Grasso JR, Hantz D, Desvarreux P, Jeannin M, Giraud A (2002) Probabilistic approach to rock fall hazard assessment: potential of historical data analysis. Natural Hazards and Earth System Sciences 2:15–26

Evans IS, McClean CJ (1995) The land surface is not unifractal: variograms, cirque scale and allometry. Z Geomorph N F Suppl. 101:127–147

Feder J (1988) Fractals. Plenum Press, New York

Frette V, Christensen K, Malthe-Sørenssen A, Feder J, Jøssang T, Meakin P (1996) Avalanche dynamics in a pile of rice. Nature 379:49–52

Fuyii Y (1969) Frequency distribution of the magnitude of landslides caused by heavy rainfall. Seismol Soc Japan J 22:244–247

Guzzetti F, Reichenbach P, Wieczorek GF (2003) Rockfall hazard and risk assessment in the Yosemite Valley, California, USA. Natural Hazards and Earth System Sciences 3:491–503

Guzzetti F, Reichenbach P, Ghigi S (2004) Rockfall hazard and risk assessment in the Nera River Valley, Umbria, Region, central Italy. Environ Manage 34:198–208

Hack JT (1957) Studies of longitudinal profiles in Virginia and Maryland. No. 294-B in US Geol. Survey Prof. Papers, US Government Printing Office, Washington D.C.

Hamon D, Nicodemi M, Jensen HJ (2002) Continuously driven OFC: A simple model of solar flare statistics. Astronomy & Astrophysics 387:326–334

Hergarten S (2002) Self-Organized Criticality in Earth Systems. Springer, Berlin, Heidelberg, New York

Hergarten S (2003) Landslides, sandpiles, and self-organized criticality. Natural Hazards and Earth System Sciences 3:505–514

Hergarten S (2012) Topography-based modeling of large rockfalls and application to hazard assessment. Geophys Res Lett 39:L13,402, DOI 10.1029/2012GL052090

Hergarten S, Krenn R (2011) Synchronization and desynchronization in the Olami-Feder-Christensen earthquake model and potential implications for real seismicity. Nonlin Processes Geophys 18:635–642, DOI 10.5194/npg-18-635-2011

Hergarten S, Neugebauer HJ (1998) Self-organized criticality in a landslide model. Geophys Res Lett 25:801–804

Hergarten S, Neugebauer HJ (2000) Self-organized criticality in two-variable models. Phys Rev E 61:2382–2385

Hergarten S, Neugebauer HJ (2001) Self-organized critical drainage networks. Phys Rev Lett 86:2689–2692

Horton RE (1945) Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Bull Geol Soc Am 56:275–370

Hovius N, Stark CP, Allen PA (1997) Sediment flux from a mountain belt derived by landslide mapping. Geology 25:231–234

Hungr O, Evans SG, Hazzard J (1999) Magnitude and frequency of rock falls along the main transportation corridors of southwestern British Columbia. Can Geotechn J 36:224–238

Jensen HJ (1998) Self-Organized Criticality – Emergent Complex Behaviour in Physical and Biological Systems. Cambridge University Press, Cambridge, New York, Melbourne

Juanico DE, Longjas A, Batac R, Monterola C (2008) Avalanche statistics of driven granular slides in a miniature mound. Geophys Res Lett 35:L19,403, DOI 10.1029/2008GL035567

Katz O, Aharonov E (2006) Landslides in vibrating sand box: What controls types of slope failure and frequency magnitude relations? Earth Planet Sci Lett 247:280–294, DOI 10.1016/j.epsl.2006.05.009

Klar AE, Aharonov E, Kalderon-Asael B, Katz O (2011) Analytical and observational relations between landslide volume and surface area. J Geophys Res 116:F02,001, DOI 10.1029/2009JF001604

Kramer S, Marder M (1992) Evolution of river networks. Phys Rev Lett 68:381–384

Kühni A, Pfiffner OA (2001) The relief of the Swiss Alps and adjacent areas and its relation to lithology and structure: topographic analysis from a 250-m DEM. Geomorphology 41:285–307

Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004) Landslide inventories and their statistical properties. Earth Surface Processes and Landforms 29:687–711

Mandelbrot BB (1967) How long is the coast of Britain? Statistical self-similarity and fractional dimension. Science 156:636–638

Mandelbrot BB (1982) The Fractal Geometry of Nature. W. H. Freeman & Co., San Francisco

Mandelbrot BB (1985) Self-affine fractals and fractal dimension. Physica Scripta 32:257–260

Noever DA (1993) Himalayan sandpiles. Phys Rev E 47:724–725

O'Callaghan JF, Mark DM (1984) The extraction of drainage networks from digital elevation data. Computer Vision, Graphics, and Image Processing 28:323–344

Olami Z, Christensen K (1992) Temporal correlations, universality, and multifractality in a spring-block model of earthquakes. Phys Rev A 46:R1720–R1723

Olami Z, Feder HJS, Christensen K (1992) Self-organized criticality in a continuous, nonconservative cellular automation modeling earthquakes. Phys Rev Lett 68:1244–1247, DOI 10.1103/PhysRevLett.68.1244

Piegari E, Cataudella V, Maio RD, Milano L, Nicodemi M (2006a) A cellular automaton for the factor of safety field in landslides. Geophys Res Lett 33:L01,403

Piegari E, Cataudella V, Maio RD, Milano L, Nicodemi M (2006b) Finite driving rate and anisotropy effects in landslide modeling. Phys Rev E 73:026,123

Piegari E, Di Maio R, Milano L (2009) Characteristic scales in landslide modelling. Nonlinear Processes in Geophysics 16:515–523, DOI 10.5194/npg-16-515-2009

von Poschinger A (2011) The Flims rockslide dam. In: Evans SG, Hermanns RL, Strom A, Scarascia-Mugnozza G (eds) Natural and artificial rockslide dams, Springer, Berlin, Heidelberg, New York, pp 407–421

Rinaldo A, Rodriguez-Iturbe I, Rigon R, Bras RL (1993) Self-organized fractal river networks. Phys Rev Lett 70:822–825

Sapozhnikov VB, Foufoula-Georgiou E (1996) Do the current landscape evolution models show self-organized criticality? Water Resour Res 32:1109–1112

Sapozhnikov VB, Foufoula-Georgiou E (1997) Experimental evidence of dynamic scaling and indications of self-organized criticality in braided rivers. Water Resour Res 33:1983–1991

Simonett DS (1967) Landslide distribution and earthquakes in the Beweni and Torricelli Mountains, New Guinea. In: Jennings JN, Mabutt JA (eds) Landform Studies from Australia and New Guinea, Cambridge University Press, Cambridge, pp 64–84

Strahler AN (1952) Hypsometric (area-altitude) analysis of erosional topography. Bull Geol Soc Am 63:1117–1142

Takayasu H, Inaoka H (1992) New type of self-organized criticality in a model of erosion. Phys Rev Lett 68:966–969

Volkwein A, Schellenberg K, Labiouse V, Agliardi F, Berger F, Bourrier F, Dorren LKA, Gerber W, Jaboyedoff M (2011) Rockfall characterisation and structural protection – a review. Nat Hazards Earth Syst Sci 11:2617–2651, DOI doi:10.5194/nhess-11-2617-2011

Voss RF (1985) Random fractal forgeries. In: Earnshaw RA (ed) Fundamental Algorithms in Computer Graphics, Springer, Berlin, Heidelberg, New York, pp 805–835

Chapter 12
SOC and Solar Flares

Paul Charbonneau

12.1 Introduction: solar magnetic activity and flares

Solar activity is made up of a broad range of physical phenomena, including geoeffective eruptive events such as flares and coronal mass ejection, as well as modulation of the Sun's corpuscular and radiative output over the full breadth of the observational energy range. Moreover, stellar observations in the era of space observatories have revealed that the myriad of phenomena collectively making up solar activity take place on every star in a comparable mass range and evolutionary stage, when observed with sufficient sensitivity.

The Sun's magnetic field is the engine for all of solar activity. Produced by a hydromagnetical dynamo process taking place at least in part in the deep interior, this magnetic field is imbedded in the strongly turbulent environment of the solar convection zone, spanning the outer 30% of the Sun's radius. The broad range of spatial scales characterizing this turbulent convection is imprinted on the solar magnetic field, which emerges at the solar photosphere in the form of structures covering at least 6 orders of magnitude in magnetic flux, their size distributed as a well-defined power law (Parnell et al. 2009). The outward extension of this photospheric flux, in turn, structures the solar corona wind, with its scale invariance in turn imprinted on a vast range of coronal and interplanetary phenomena.

This chapter focuses on one specific solar phenomenon, flares, for which a very strong physical case for a SOC scenario can be made. Solar flares, most readily observed in the short-wavelength range of the electromagnetic spectrum (extreme ultraviolet to soft X-ray, from about 100nm down to 1nm), result from the impulsive and spatially localized release of energy in the solar corona (see, e.g., Benz 2008, and references therein). Figure 12.1 shows one particularly spectacular example, having occurred on 04 November 2003. The image was taken in the extreme ultraviolet (19.5nm) by the EIT instrument onboard the SOHO spacecraft. This flare, occurring over a large active

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

region (AR10486 in the NOAA numbering convention) just about to rotate off the visible solar disk, remains to this day the most intense X-ray flare ever recorded from space. The bottom panel on Fig. 12.1 shows X-ray time series covering that same time period. The flare of Fig. 12.1 corresponds to the tallest peak, at UT 19:48 on Nov 4. Note also the many other large flares having taken place in the preceeding 48 hours, half of which originating over active region AR10486. The many smaller X-ray spikes peppering the time series are also flares, of much smaller magnitudes, many occurring elsewhere on the solar disk, and blending into the more slowly varying background level arising from diffuse emission over the whole Sun.

The vast majority of large flares, including that on Fig. 12.1, occur within large solar active regions, which are strongly magnetized complexes coincident at photospheric levels with sunspots and other strongly magnetized structures. Even the smaller flares tend to occur in regions overlying magnetic structures, sometimes much smaller than bona fide active regions. This spatial coincidence leaves little doubt that the solar magnetic field provides the energy reservoir powering the flaring phenomenon. Hard X-ray observations (see, e.g., Masuda et al. 1994) indicate that flare onset usually takes place above the photosphere, in the lower corona, where the electrical conductivity is quite large and therefore the magnetic diffusion time many orders of magnitude longer than observationally-inferred onset times, which are of the order of seconds (viz. Fig. 12.1). This extreme disparity between observed onset time and the magnetic diffusion timescale leaves magnetic reconnection as the most likely (if not the only) physical candidate for a dynamical (rather than diffusive) release of energy (see Shibata & Magara 2011, and references therein). Despite decades of research on this topic, the instability(ies) responsible for the triggering of magnetic reconnection have not yet been convincingly identified, although numerous candidate magnetohydrodynamical and plasma instabilities have been proposed.

Flares have been extensively studied, in part because of their potentially damaging impact on communications networks and technological infrastructures, especially on satellites in high orbits where the shielding effect of Earth's magnetic field is reduced. From the very largest flares, so intense as to damage satellite electronics, solar panels and imaging instruments and even pose a health threat to orbiting astronauts, flare-like energy release events have been observed down to the instrumental detection limit. The inferred frequency distribution of flare energy takes the form of a well-defined power-law, all in all covering over eight orders of magnitude in flare energy (Dennis 1985; Aschwanden et al. 2000; Aschwanden 2011). While the conversion of observed EUV or X-rays fluxes to flare energy does involve a number of working hypotheses regarding the geometry of the flaring volume, physical conditions therein, and physical approximation in the treatment of hard radiation emission (e.g., McIntosh & Charbonneau 2001; Aschwanden & Parnell 2002, McIntosh 2000; Parenti et al. 2006; and references therein), solar flares stand as one of the most impressive examples of scale-invariance to be

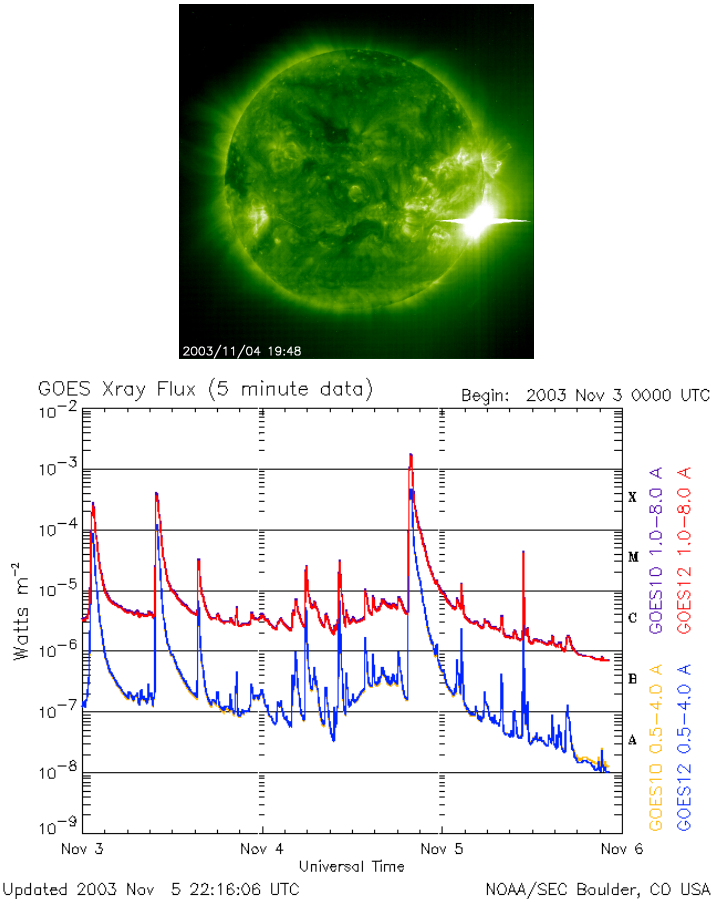


Fig. 12.1: Top: Extreme Ultraviolet (19.5nm) image of the solar disk on 4 Nov 2003, capturing a large flare near its peak phase (lower right quadrant; the two thin, bright horizontal spikes are artefacts caused by oversaturation of pixels in the camera's CCD). This flare is the most intense X-ray flare ever recorded from space, and topped off a series of strong flares, triggered over active region AR10486 and neighbouring AR10488, starting on 28 October with the so-called "Halloween storm" event. Image obtained by the EIT intrument onboard SOHO (Solar and Heliospheric Observatory), a joint mission of NASA and ESA. Bottom: a time series of GOES soft-X-rays fluxes in two wavelength bands, as color-coded, over a 3-day time period including the large flare on the above EIT/SOHO image, late on 4 Nov. Note the very rapid onset phase (the vertical scale is logarithmic!), followed by a slower decay phase, a pattern largely independent of flare size. Plot courtesy of NOAA/SEC.

encountered in astrophysics. As remarkable perhaps, the power-law slope of the PDF of flare energy release remains invariant in the course of the solar cycle (Lu et al. 1993; Aschwanden 2011; Aschwanden & Freeland 2012), even though flaring rates vary by many orders of magnitude between cycle minima and maxima.

Understanding and modelling the flaring process presents a number of substantial challenges beyond those posed by the ill-understood physics of magnetic reconnection in the solar corona; a large flare, such as on Fig. 12.1, can liberate a few 10^{33} erg of energy in a few minutes, which is approaching 1% of the per-second solar radiative luminosity. Moreover, the same active region can produce many substantial flares within a relatively short period of time (viz. Fig. 12.1). This apparent lack of "recovery time" indicates that even a very large flare does not necessarily deplete its energy reservoir to a significant extent. While many solar active regions exhibit higher flare activity during and immediately after their emergence at the photosphere, large flares can also occur in mature active regions undergoing little no magnetic evolution, at least on the scales resolved by photospheric magnetograms. In other words, flaring often appears "spontaneous", in the sense of having no obvious observable external trigger, analogous in this sense to magnetospheric substorms. All this suggests that flares are not directly driven by subphotospheric processes such as magnetic flux emergence, but arise naturally as part of intrinsic coronal dynamics.

The primary question is thus: how can the vast amounts of energy required to power the flaring phenomenon be pumped into the corona, temporarily stored there, and impulsively released in the observed scale-invariant manner? It turns out that self-organized criticality can provide a viable answer. However, the path towards this answer runs through what will look at first like a conceptual detour —but will turn out not to be— involving another great solar contemporary mystery, namely coronal heating.

12.2 Parker's coronal heating hypothesis

The corona is the Sun's outer atmosphere, and extends many solar radii into interplanetary space where it smoothly blends into the expanding solar wind. Even away from flaring regions, its temperature reaches 1.5 million degrees kelvin, very much higher than the $\sim 6000\,\mathrm{K}$ of the photosphere. The coronal gas is very tenuous, so energetically-speaking it should not be that hard to heat it up to $10^6\,\mathrm{K}$; however, while the corona is traversed by a huge radiative energy flux, heat flows from hot to cold, so nothing like equilibrium thermodynamics will do the trick. To make matters worse, the corona is losing energy at a substantial rate. Given radiative and conductive losses at its base, and thermal losses due to acceleration of the solar wind, the average energy flux required to maintain a quasi-steady corona adds up to $\sim 10^6\,\mathrm{erg\ cm^2}$

s^{-1}, for a total of 6×10^{28} erg s^{-1} integrated over the solar photosphere (see Aschwanden 2004, §9.1). While this represents a mere $\simeq 10^{-6}$ fraction of the Sun's radiative luminosity, how the required energy flux gets diverted into the corona is not yet understood quantitatively (see, e.g., Klimchuk 2006 for a review).

One very potent energy reservoir available to heat the corona is the mechanical energy associated with subphotospheric thermal convection, but again the challenge is to transfer energy from that reservoir into the corona. Copious emission of acoustic and magnetosonic waves is expected at photospheric levels, but these are very rapidly absorbed as they propagate upwards, contributing mostly to the heating of the chromosphere, a thin, 10^4 K intermediate atmospheric layer sandwiched between the photosphere and the base of the corona. Alfvén waves face the opposite problem: in magnetically open regions they readily propagate up and through the corona, suffering very little absorption and escaping instead into interplanetary space.

It turns out that the coronal magnetic field, ultimately anchored in these same subphotospheric convective layers, can act as a transport and storage agent for this mechanical energy. The idea has a long and tortuous history, but was worked out in detail in the 1980's by E.N. Parker (see Parker 1983, 1988), and is illustrated in cartoon form on Figure 12.2.

It is an observed fact that the solar corona is filled with magnetic loops, anchored at both ends on unipolar magnetic flux concentrations of opposite polarity scattered all over the photophere. Many such loops are actually visible over the solar limb above the flare on Fig. 12.1, and can extend over substantial heights well into the corona. Because of the high electrical conductivity of the coronal plasma, a very strong dynamical coupling exist between the plasma and magnetic field. The latter is said to be "frozen" in the plasma.

Imagine now an "unbent" coronal loop, where the photospheric footpoints have been cut out and the loop turned into a rectilinear magnetic structure anchored on two planes at "top" and "bottom", as depicted on Figure 12.2A. At and below the photosphere, the energy density of the plasma greatly exceeds the energy density of the magnetic field in a typical coronal loop. As a result, even if the loop was originally made of purely parallel magnetic fieldlines, as drawn on Fig. 12.2A, their footpoints will be randomly shuffled by the horizontal convective fluid motions at photospheric levels. In the corona, on the other hand, due to the sharp drop in density with height (from an few 10^{-7} g cm^{-3} at photospheric levels, down to $\sim 10^{-16}$ g cm^{-3} at the coronal base, 3000 km higher up), it is the energy density of the magnetic field that now largely exceeds that of the plasma. The coronal magnetic field will consequently attempt to relax to a force-free configuration in response to the footpoint displacements, but the high conductivity of the coronal plasma poses a strong topological constraint: magnetic fieldlines cannot cross! As a consequence, the magnetic field becomes entangled (A→C on Figure 12.2), reaching a quasi-static force-free state everywhere except where fieldlines are forced to wrap around one another, forming a local tangential discontinuity

with which is associated a current sheet within which is flowing a strong electrical current density ($\mathbf{J}$), as per Ampère's Law:

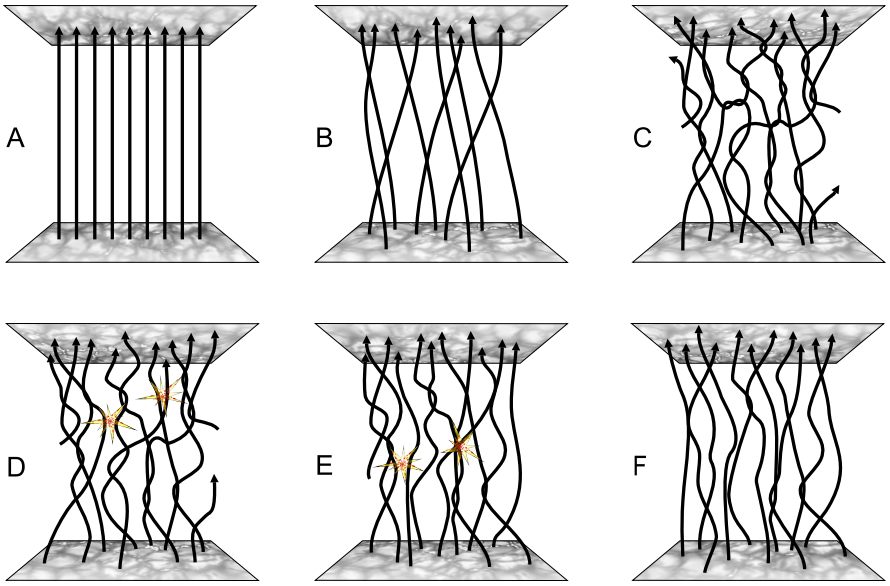


Fig. 12.2: Cartoon representation of an "unbent" coronal loop. The top and bottom planes both represent the solar photosphere, where horizontal convective motions inexorably shuffle the fieldlines' footpoints (A→B→C). Magnetic tangential discontinuities form as the loop attempts to relax to a force-free state under the topological constraints dictated by the high electrical conductivity (C). Instability of these tangential discontinuities leads to magnetic reconnection, spatially localized energy release (D), and disruption of neighbouring marginally stable current sheets (D→E). While the cascade of reconnection events has reduced the overall stress (cf. panels C and F), the system nonetheless retains a higher level of internal magnetic stress than in its original configuration (panel A). Graphics courtesy of D. Passos.

$$\frac{4\pi\mathbf{J}}{c} = \nabla \times \mathbf{B} \tag{12.1}$$

(CGS units are used throughout). As the magnitude of this current keeps growing, it can become unstable with respect to a number of plasma instabilities, leading to reconnection and rapid, spatially localized release of magnetic energy at the sites of these tangential discontinuities (Fig. 12.2D).

Parker (1988) presents a simple order-of-magnitude estimate of the associated energy input into the corona, which is well worth recalling here. Consider a coronal loop of length L and magnetic field strength B, with footpoints an-

chored vertically into the photosphere and subjected there to advection of an horizontal flow of magnitude v and typical length scale ℓ. Over a time t this flow will induce an angular deviation $\theta(t)$ with respect to the vertical of magnitude given by

$$\tan\theta(t) \simeq vt/L \ . \tag{12.2}$$

The associated horizontal magnetic component B_h has thus a strength

$$B_h = B\,\tan\theta(t) \simeq \frac{Bvt}{L} \ . \tag{12.3}$$

Now, in the high electrical conductivity regime of relevance here, a magnetic field acted upon in this manner by a flow will resist the flow through the buildup of a magnetic stress (Lorentz force per unit area) with horizontal component of order $S \sim B_h B/4\pi$; the work done per unit surface by the horizontal flow against this force in displacing a footpoint over an infinitesimal distance $\mathrm{d}x$ is then

$$\mathrm{d}W = \frac{B_h B}{4\pi}\mathrm{d}x \ , \tag{12.4}$$

so that, making use of eq. (12.3):

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \frac{vB_hB}{4\pi} \simeq \frac{B^2v^2t}{4\pi L} \ , \qquad [\mathrm{erg\ cm^{-2}\ s^{-1}}] \tag{12.5}$$

since $v = \mathrm{d}x/\mathrm{d}t$. Parker goes on to argue that for plausible coronal loop parameters ($B \sim 100\,\mathrm{G}$, $v \sim 0.5\,\mathrm{km\ s^{-1}}$, $L \sim 10^{10}\,\mathrm{cm}$), a heating rate of 10^7 erg cm^2 s^{-1}, corresponding to total coronal energy losses over active regions, is reached in a time interval $t \sim 5 \times 10^4\,\mathrm{s}$, amounting to a maximal deflection angle of only $\theta_c \sim 14°$. Now, this limiting angle, corresponding to the angle subtented by pairs of magnetic fieldlines bending around one another on Fig. 12.2, is likely set by whichever instability is responsible for triggering magnetic reconnection; Parker observes that a higher instability threshold θ_c would lead to higher B_h, and thus higher $\mathrm{d}W$. In other words, the *global* volumetric heat input into varies inversely with the efficiency of *local* dissipative mechanism.

Only one additional "ingredient" needs to be added to this physical picture to produce a model for flares of all sizes: the possibility that reconnection at one unstable site alters the physical conditions at neighbouring sites of tangential discontinuities in a manner such as to sometimes push these sites over their stability threshold (D→E on Figure 12.2). Reconnection at these sites can then trigger further reconnections at other neighbouring sites, and so on in classical avalanching style until stability is restored across the whole structure (Fig. 12.2F). Under this view, a large flare is just an ensemble of elementary reconnection events (dubbed "nanoflares") cascading through a

stressed coronal loop. Larger (smaller) flares simply involve the avalanching of more (fewer) elementary reconnection events, which is good recipe for scale invariance. Note also that, as per the above power estimate, a (relatively small) coronal loop of photospheric cross-section $1000 \times 1000\,\mathrm{km}^2$ (comparable to a granule) can accumulate $\sim 10^{28}$ erg of energy in about 10 hours, which is comparable to the energy output of a mid-size flare.

Perhaps more importantly in the context of this chapter, this modified form of Parker's coronal heating conjecture contains all the components believed to be conducive to classical self-organized critical behavior (see, e.g., Jensen 1998): an open dissipative system (a coronal loop) undergoing slow forcing (by photospheric fluid motions) and subject to a self-limiting local threshold instability (magnetic reconnection).

12.3 SOC Models of solar flares

In retrospect, it was perhaps inevitable that following the seminal SOC papers of Bak, Tang & Wiesenfeld (1987, 1988; hereafter BTW), someone would seek to interpret the remarkable scale invariance observed in solar flare statistics in terms of a SOC process. Inspired by BTW but following more closely Kadanoff et al. (1989), Lu & Hamilton (1991; hereafter LH91) developed the first SOC sandpile model targeted at solar flares, that by now has become a kind of "standard". The many variations on the theme set by LH91 that soon followed are reviewed in Charbonneau et al. (2001), to which we refer the interested reader. Nearly twenty years later it should come as no surprise that such models can be designed in such as way as to reproduce quite well the observed power-law form of flare parameter statistics (peak flux, total released energy, duration, etc.), including the numerical value of the associated exponents. The real challenge remains at the level of the physical interpretation to be attached to the various sandpile evolutionary rules. Accordingly, in this section we describe the workings of the basic LH91 model (or, more precisely, the variant subsequently published by Lu et al. (1993), hereafter LH93; see also Lu 1995a and Vlahos et al. 1995). In the following section we delve into the specific physical interpretations having been tagged onto this model.

12.3.1 The Lu & Hamilton model

The LH91 and LH93 models are cellular automata defined over a 3D regular cartesian grid with nearest-neighbour connectivity (top+ down+ right+ left+ front+ back) over which a vector field $\mathbf{A}$ is defined. Here we will consider instead a scalar version of the model, where a nodal quantity $A^t_{i,j,k}$ is defined

over the lattice[1]. For the time being we will simply consider that the nodal variable A is related in some way to the magnetic field. The superscript t is a discrete time index, and the subscript triad (i, j, k) identifies a single node on the 3D lattice.

Keeping $A = 0$ on the lattice boundaries, the cellular automaton is driven by adding one small increment of A per time step, at some randomly selected node that changes from one time step to the next:

$$A_{\mathbf{r}}^{t+1} = A_{\mathbf{r}}^{t} + \delta A \ , \tag{12.6}$$

where the vector index $\mathbf{r}$ denotes a random node (i, j, k) on the 3D lattice. The increment δA is extracted from a random distribution with non-zero mean, so that its action over a large number of time steps will lead to the buildup of A over the whole lattice, much as in any classical sandpile model.

A deterministic stability criterion is now defined in terms of the local curvature of the field at node (i, j, k):

$$\Delta A_{i,j,k}^{t} = A_{i,j,k}^{t} - \frac{1}{6} \sum_{\mathbf{n}} A_{\mathbf{n}}^{t} \ , \tag{12.7}$$

where the sum runs over the six nearest neighbours at nodes $(i, j, k \pm 1)$, $(i, j \pm 1, k)$ and $(i \pm 1, j, k)$, hereafter denoted by the vector index $\mathbf{n}$. If this quantity exceeds some pre-set threshold Z_c:

$$|\Delta A_{i,j,k}^{t}| > Z_c \ , \tag{12.8}$$

then an amount of nodal variable Z is redistributed to the same set of nearest neighbours according to the following discrete, conservative deterministic rules:

$$A_{i,j,k}^{t+1} = A_{i,j,k}^{t} - \frac{6}{7} Z \ , \tag{12.9}$$

$$A_{\mathbf{n}}^{t+1} = A_{\mathbf{n}}^{t} + \frac{1}{7} Z \ , \tag{12.10}$$

with $Z = \Delta A_{i,j,k}^{t}$ in the Lu & Hamilton (1991) model, but $Z = \mathrm{sign}(\Delta A_{i,j,k}^{t}) \times Z_c$ in the LH93 model considered here[2]. Following this redistribution it is possible that one of the nearest-neighbour nodes now exceeds the stability threshold. The redistribution process begins anew from this node, and so on in classical avalanching manner. Driving is suspended during avalanching, implicitly implying a separation of timescales between driving and avalanching dynam-

[1] Robinson (1994) has shown that for the type of driving used in the Lu & Hamilton model, the use of a vector nodal variable yields results identical (statistically) to the use of a scalar variable.

[2] For slow driving ($|\delta A| \ll A$), the stability criterion (12.7) will (typically) only be slightly exceeded when the instability is triggered, so these two choices yield similar evolution, although their numerical implementation exhibit distinct stability properties; see Liu et al. (2003) for more on these matters.

ics ("stop-and-go" sandpile), and all nodal values are updated synchronously during avalanche to avoid introducing a directional bias in avalanche propagation, after which the $A = 0$ condition is enforced at boundary nodes before proceeding with the subsequent time iteration.

It is readily shown that these redistribution rules, while conservative in A, lead to a decrease in A^2 summed over the seven nodes involved in the redistribution by an amount:

$$\Delta e^t_{i,j,k} = \frac{6}{7}\left(2\frac{|\Delta A^t_{i,j,k}|}{Z_c} - 1\right) Z_c^2 \ , \tag{12.11}$$

with the energy released being "assigned" to the unstable node (i, j, k). If one identifies A^2 with a measure of magnetic energy (more on this shortly), the total energy liberated by all unstable nodes at a given iteration is then equated to the energy release per unit time in the flare:

$$(\Delta E)^t = \sum_{\text{unstable}} \Delta e^t_{i,j,k} \ , \tag{12.12}$$

and the lattice energy at iteration t is simply given by

$$E^t = \sum_{\text{all nodes}} (A^t_{i,j,k})^2 \ . \tag{12.13}$$

Note that strictly speaking, the expected relation

$$E^{t+1} = E^t - (\Delta E)^t \ , \tag{12.14}$$

only holds for avalanches that do not reach the lattice boundaries, since the imposed boundary conditions $A = 0$ also remove energy from the system in a manner not accounted for in eq. (12.12). In practice, eq. (12.14) is found to hold very well except on very small lattices.

A natural energy unit here, used in all that follows, is the quantity of energy e_0 liberated by a single node exceeding the stability threshold by an infinitesimal amount; setting $\Delta A_{i,j,k} \simeq Z_c$ in eq. (12.7) immediately leads to

$$e_0 = (6/7)Z_c^2 \ . \tag{12.15}$$

This is the model's analog of nanoflares in Parker's scenario previously discussed.

12.3.2 Sample results

Figure 12.3, taken directly from Charbonneau et al. (2001) for illustrative purposes, shows time series of lattice "energy" (eq. 12.13) and released energy (eq. 12.12), produced by a 2D version of the LH93 model run on a small 48×48 lattice, starting from an empty lattice, i.e., $A^0_{i,j,k} = 0$ at all nodes. As the lattice energy gradually builds up, progressively larger avalanches are triggered intermittently. After about 12 million iterations, the lattice energy abruptly levels off, coinciding with a marked increase in the size of the largest avalanches cascading over the lattice. At this point in time the SOC state has been reached, A remains thereafter in a statistically stationary state characterized by what looks like constant curvature over the global scale of the lattice, and avalanche of sizes going from a single node up to the lattice size evacuate A at the boundaries nodes (where $A = 0$ is enforced at every time iteration) at mean rate equal to that associated with the random driving process. As shown already in LH93, this very simple model yields a good representation of flare statistics, namely the observed power-law form (and associated exponents) of the frequency distributions of flare peak energy release, total energy release, and, to a lesser extent, duration.

Figure 12.4 shows the distribution of the nodal variable A (top) and curvature measure ΔA scaled to the threshold value Z_c, over the 2D lattice in the SOC state at the end of the same simulation as on the previous Figure. The SOC state appears quite smooth on the scale of the top plot, yet in fact the curvature measure (bottom) shows a lot of structure down to the scale of the lattice, as one would have expected from the spatiotemporally discrete and stochastic nature of the driving process used here. Indeed, even in the statistically stationary SOC state the global curvature characterizing the top plot is about 70% of the curvature corresponding to the instability threshold Z_c, meaning that the curvature at most nodes is significantly below threshold. At this point in this simulation, a large avalanche is in the process of dying off, with 10 nodes still exceeding threshold here. These nodes are grouped in three small clusters, the most prominent near lattice center. As with other sandpiles using stability criteria based directly on nodal value, or its slope, it is quite common for large avalanches to be rather compact in their growth phase, but fragment into spatially distinct sub-avalanches later on. The automated identification of avalanches is not a trivial task here, as they must be treated as spatiotemporal structures (see, e.g., Uritsky et al. 2007).

Figure 12.5 shows the ensemble of nodes having avalanched at least once in the course of a large avalanche on a 3D version of the LH93 model ran on a $64 \times 64 \times 64$ lattice. Geometrical analyses of many such avalanches confirms their fractal nature, with dimension 1.8 for the original LH93 model in 3D at peak time, and 1.6 for the 2D version. An interesting exercise consists in projecting this 3D fractal structure on a 2D observational plane, where each pixel in this plane is assigned an intensity equal to the total energy released

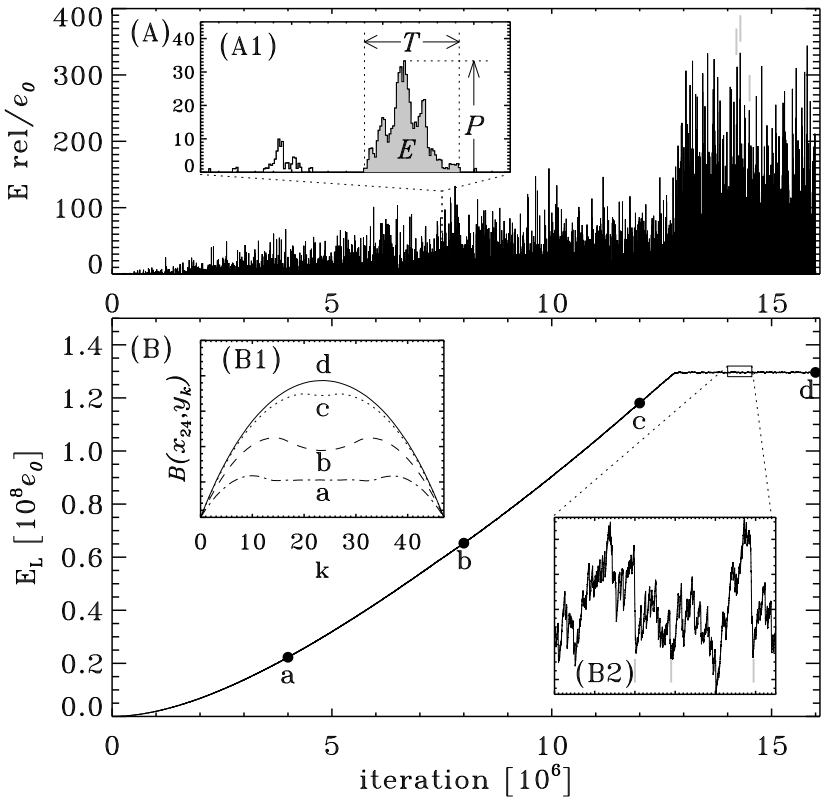


Fig. 12.3: Time series of energy released by avalanches (A), as given by eq. (12.12), and total lattice energy (B), as given by eq. (12.13), in a 2D scalar version of the LH93 model, on a very small lattice of size 48×48. Note how the transition to a statistically stationary SOC state, with lattice energy leveling off in B, is accompanied by a marked increase of the largest avalanches running across the lattice, in A. Inset A1 is a closeup showing the structure of a few individual avalanches. Inset B1 shows cuts across the lattice center at the four epochs indicated by solid dot on the time series in B, as labeled. Inset B2 is a close up on the lattice energy time series in the SOC state, sudden drops corresponding to avalanches. Taken from Charbonneau et al. (2001), Figure 3.

on the corresponding line-of-sight through the avalanche (a procedure appropriate for an optically-thin medium). Three such synthetic "flare images"

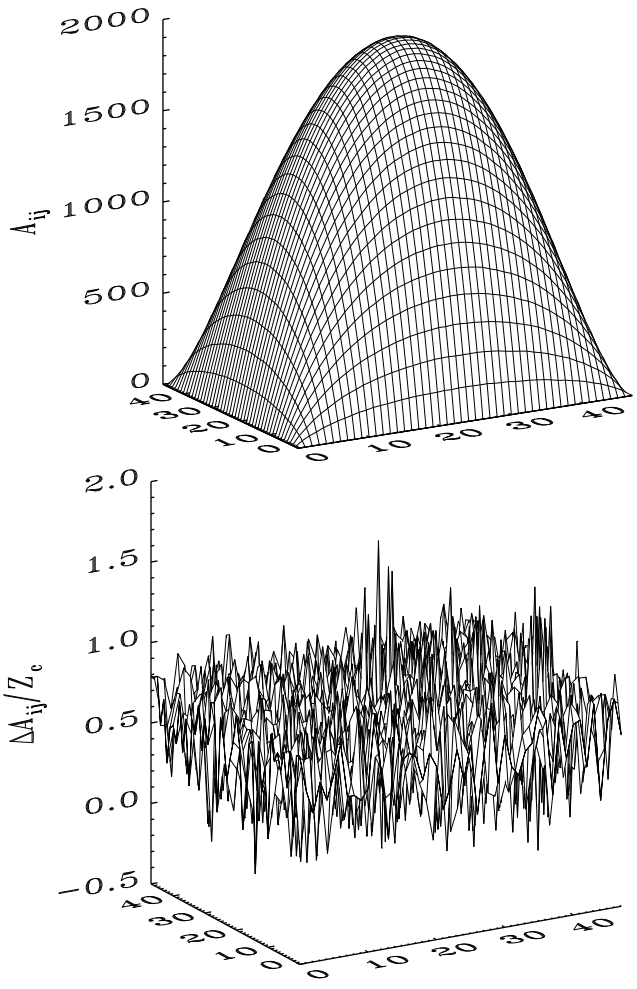


Fig. 12.4: Snapshot of the spatial profiles of the nodal variable $A_{i,j}$ (top) and curvature measure $\Delta A_{i,j}$ (bottom), the latter normalized to the stability threshold $Z_c = 5$, as extracted at the last iteration of the 48×48 2D LH93 simulation of Figure 12.3. While the nodal variable presents what looks like a smooth distribution over the lattice on the scale of this plot, the curvature measure shows a lot of node-to-node fluctuations about a mean value correponding to the mean curvature so obvious on the top plot.

are plotted at right on Fig. 12.4, each corresponding to a different projection plane. These projections are also fractal structures.

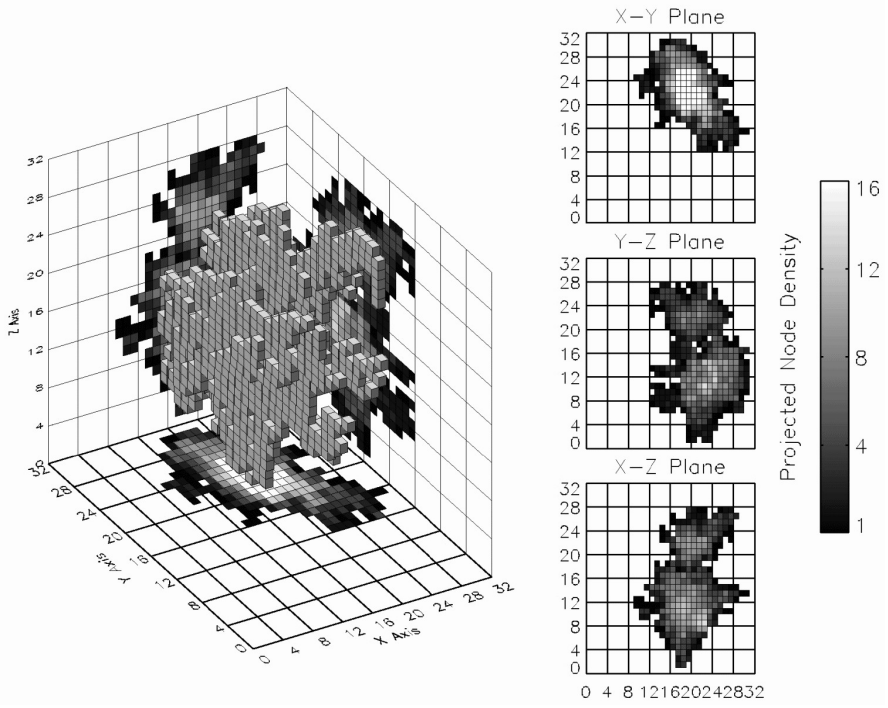


Fig. 12.5: A large avalanche in the LH93 model on a $64 \times 64 \times 64$ lattice. A small cube is drawn at each lattice node having gone unstable at least once in the course of this avalanche. The resulting structure is a fractal. The three panels on the right show "views" of this avalanche from the three cartesian planes defined by the lattice, where each pixel is assigned a grayscale intensity proportional to the amount of energy liberated by all nodes located along the corresponding line-of-sight through the course of the avalanche, as would be observed by an "imaging instrument" observing the avalanche at a time cadence much longer than its evolutionary timescale. These projected "images" are also fractal. Taken from McIntosh et al. (2002), Figure 5.

12.4 Physical interpretation

For all its success at reproducing the statistical distributions of flare parameters, the discrete CA rules governing the operation of the LH93 model remain very far removed from the MHD equations presumably governing the process of magnetic reconnection and dynamical reconfiguration of the coronal magnetic field. In this section we review in some detail the physical interpretation that offers arguably the most convincing link to coronal magnetohydrodynamics. The task is to assign physically well-grounded meaning to (1) the lattice; (2) the nodal variable; (3) the random forcing; (4) the sta-

bility criterion; (5) the calculation of energy released in avalanches; and (6) the redistribution rules. We consider these various components in turn.

12.4.1 The lattice and nodal variable

We adopt here a variation of the interpretative picture originally put forth by Lu et al. (1993) and much elaborated by others (e.g., Vassiliadis et al. 1998; Isliker et al. 1998, 2000, 2001). The 3D lattice will be considered as the interior of a an "unbent" coronal loop (viz. Fig. 12.2), and the nodal variable A as the z-component of the magnetic vector potential $\mathbf{A}$, such that

$$\mathbf{B} = \nabla \times \mathbf{A} \ . \tag{12.16}$$

Note already that any magnetic field defined in this manner will automatically satisfies the solenoidal constraint $\nabla \cdot \mathbf{B} = 0$. The vector potential component A_z then defines the magnetic field component in the xy plane, which can be directly related to the degree of twist in the loop, i.e., the ratio of transversal to longitudinal field components[3]. As argued by Lu et al. (1993), adding small random increments of A at one lattice node then corresponds to adding a small amount of twist somewhere in the loop, which fits nicely the Parker scenario outlined above.

Isliker et al (2001) show that under this *Ansatz*, a number of realistic magnetic configurations can be produced from the original LH93 model (with the nodal variable a vector quantity, namely the full magnetic vector potential $\mathbf{A}$). Under different boundary conditions and driving schemes (isotropic vs anisotropic), the magnetic field in the SOC state can take the form of solenoidal structures organized around magnetic neutral lines, or arcades of magnetic fieldlines connected to a boundary plane (see Isliker et al. 2001, Figs. 1 and 2).

12.4.2 The stability criterion

Consider now Ampère's Law (12.1), again written in terms of the vector potential:

$$\mu_0 \mathbf{J} = \nabla \times (\nabla \times \mathbf{A}) = -\nabla^2 \mathbf{A} \ , \tag{12.17}$$

under the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$. Next, note that through simple algebraic manipulation the curvature measure (12.7) can be rewritten as

[3] The longitudinal magnetic field is assumed to remain unaffected by the variations of the field components contained within any plane perpendicular to the loop axis, as required by conservation of magnetic flux along the loop.

$$\Delta A^t_{i,j,k} = -\frac{\Delta^2}{6}\Big(\frac{A_{i+1,j,k} - 2A_{i,j,k} + A_{i-1,j,k}}{\Delta^2}$$

$$+\frac{A_{i,j+1,k} - 2A_{i,j,k} + A_{i,j-1,k}}{\Delta^2} + \frac{A_{i,j,k+1} - 2A_{i,j,k} + A_{i,j,k-1}}{\Delta^2}\Big) \ . \quad (12.18)$$

With Δ denoting the internodal nearest-neighbour distance, the quantity within parentheses on the RHS has the form of a second-order centered finite difference for the standard Laplacian operator acting on A, so that the instability criterion becomes:

$$|\nabla^2 A| > \frac{6Z_c}{\Delta^2} \ , \quad (12.19)$$

This is a threshold condition on $\nabla^2(A_z\hat{\mathbf{z}})$, i.e., the magnitude of the longitudinal (J_z) component of the electric current, as per eq. (12.17). This is quite satisfying from the point of view of MHD and plasma stability. Returning to the bottom plot on Fig. 12.4, it is also clear that the current density shows a lot of fine-scale structure, (see also Isliker et al. 2001, Fig. 6, for a 3D equivalent), which is consistent with Parker's picture of small current sheets being ubiquitous throughout coronal magnetic loops.

12.4.3 Computing the released energy

As an *Ansatz*, the identification of the nodal variable in the LH93 model with the magnetic vector potential suffers from one significant difficulty: there is not necessarily a one-to-one correspondence between the squared nodal variable and magnetic energy as conventionally defined:

$$\mathcal{E}_B = \frac{1}{8\pi}\int \mathbf{B}^2 \mathrm{d}V \equiv \frac{1}{8\pi}\int (\nabla\times\mathbf{A})^2 \mathrm{d}V \neq \frac{1}{8\pi}\int \mathbf{A}^2 \mathrm{d}V \ . \quad (12.20)$$

In the case of the 2D version of the LH93 model, with the nodal variable identified with the z-component of the magnetic vector potential, eq. (12.13) should be replaced by:

$$E^t = \sum_{\text{all nodes}} \left[\left(\frac{A^t_{i,j+1} - A^t_{i,j-1}}{2\Delta} \right)^2 + \left(\frac{A^t_{i-1,j} - A^t_{i+1,j}}{2\Delta} \right)^2 \right] \ . \quad (12.21)$$

assuming unit lattice spacing Δ and using second-order centered finite differences for the spatial derivatives. Evidently, for a fixed field geometry, one expects the overall magnitude of $\mathbf{B}$ to be proportional to that of $\mathbf{A}$; however, the energy released in the LH93 model arises primarily from the small-scale fluctuations about the mean global state (viz. Fig. 12.4 herein). We already remarked that the LH93 redistribution rules lead to a decrease of A^2 over the

nodes involved in the redistributions, as per eq. (12.11). The key question is then: are the redistribution rules really reducing the magnetic energy, as defined physically though eq. (12.21)?

Figure 12.6 offers an answer to this question. Panels (A) and (B) show energy release and lattice energy time series spanning the last 20000 iterations of the simulation run plotted on Fig. 12.3. As expected, there is a perfect temporal correlation between the occurrence or large avalanches in (A), and drops of lattice energy in (B). Panel (C) is the lattice energy time series computed this time using eq. (12.21). The overall correlation with the time series in (B) appears quite good, in that most large drops in both time series correlate temporally very well with one another; careful comparison of the two time series reveals exceptions, however, the most prominent being flagged by a vertical dashed line. The large avalanche beginning around iteration 11000 shows up as a large fluctuations in (C), but does not decrease the magnetic energy as other avalanches of similar size do. Nonetheless, as a statistical rule, avalanches do lead to a decrease of "true" magnetic energy in this model, a rule holding especially well for the larger avalanches.

To sum up, the apparent ambiguity in the definition of magnetic energy does not turn out to be problematic, and the identification of the sandpile nodal variable with the magnetic vector potential is still holding up well at this stage.

12.4.4 Nodal redistribution as nonlinear diffusion

The LH93 redistribution rules (12.9)–(12.10) clearly diperse locally the nodal variable in a manner isotropic and reducing local gradients between the unstable node and its neighbours. As such it behaves as a diffusion process, albeit a strongly nonlinear one since it is activated only once a node exceeds the stability threshold, and is nil otherwise. Lu (1995b) investigated the SOC-like avalanching behavior in a model 1D nonlinear diffusion equation of the form:

$$\frac{\partial A(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(D(A,t)\frac{\partial A(x,t)}{\partial x}\right) + S(x,t) \ , \tag{12.22}$$

where $S(x,t)$ is a spatiotemporally random driving function with non-zero mean, and the space- and time-variable diffusion coefficient D evolves continuously between floor and ceiling values $D_{\min}$, $D_{\max}$ according to:

$$\frac{\partial D}{\partial t} = \frac{Q(A)}{\tau} - \frac{D}{\tau} \ , \tag{12.23}$$

with the switching function $Q(A)$ defined in terms of the nodal variable slope through the threshold relation:

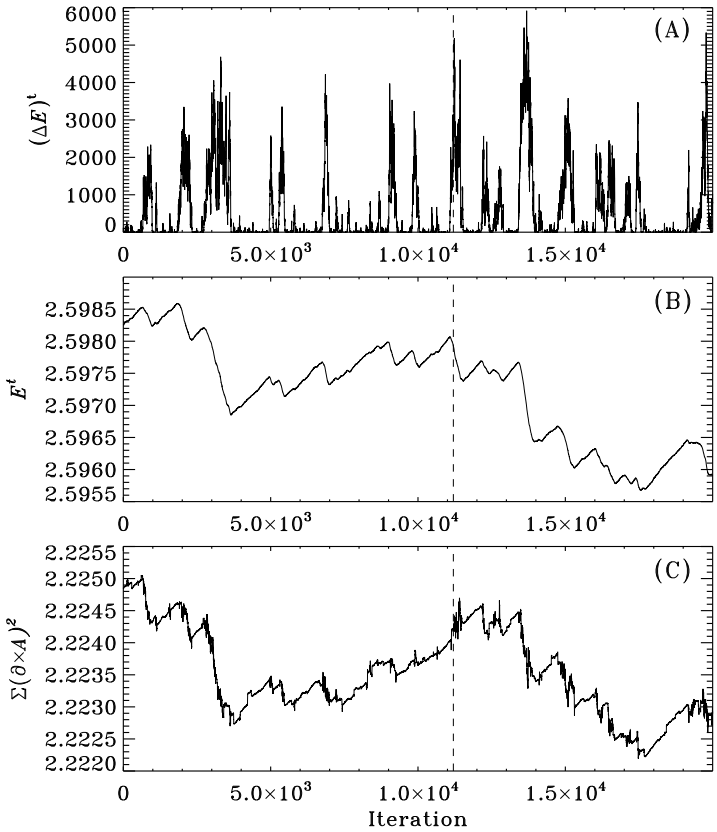


Fig. 12.6: Time series of (A) released energy, as given by eq. (12.12), and (B) lattice energy as given by eq. (12.13), correponding to the last 20000 iterations of the LH93 2D simulation of Fig. 12.3. Panel (C) is an alternate time series of magnetic energy, computed this time using eq. (12.21). Although the latter shows more fine structure, the overall correlation with the time series in (B) is excellent, especially with regards to the energy drops associated with large avalanches. The vertical dashed line identifies a large avalanche where the expected energy drop does not appear on panel C; such occurrences are the exception rather than the rule (see text). Here as in panels A1 and B2 of Fig. 12.3, in plotting the time series the same time step size is used during and in between avalanches, in order to illustrate the latter's temporal structure; under the separation of timescale characterizing stop-and-go sandpiles, avalanches and associated lattice energy drops should be nearly instantaneous in comparison to the inter-avalanche wait time (cf. Fig. 2 in Lu et al. 1993)

$$Q(A) = \begin{cases} D_{\min} & |\partial A/\partial x| < k \\ D_{\max} & |\partial A/\partial x| > \beta k \end{cases} \tag{12.24}$$

with a decay constant τ, threshold parameter k and hysterisis parameter β. Note that under these definitions, and with $0 < \beta < 1$, the gradient required to turn the instability on ($D_{\min} \rightarrow D_{\max}$) is larger than that required for the instability to sustain itself once activated. Once the local gradient has fallen below this threshold, the diffusion coefficient decays back to its floor value $D_{\min}$ on a timescale given by τ. Lu (1995b) shows that this (relatively) simple nonlinear diffusion system, when driven long enough, settles in a state where energy dissipation occurs through discrete avalanching-event with sizes distributed as a power-law, in a manner qualitatively similar to sandpile models.

12.4.5 Reverse engineering of discrete redistribution rules

The nonlinear diffusion interpretation suggested by Lu (1995b) can be formalized by considering the redistribution rules as discretized forms of differential operators, which then allows to "reverse engineer" partial differential equation describing the phenomenon, which can (perhaps) take us one step closer to true magnetohydrodynamics. In the solar flare context many such PDEs have been "inverted" from the LH93 and related cellular automata (e.g. Vassiliadis et al. 1998; Isliker et al. 1998; Liu et al. 2003). For illustrative purposes, we only describe here the reverse engineering procedure proposed in Liu et al. (2003).

For simplicity, consider a 1D version of the LH91 model, with stability criterion and redistribution rules given by:

$$\Delta A_i^t \equiv A_i^t - \frac{1}{2} \sum_{j=i\pm1} A_j^t \ , \tag{12.25}$$

$$A_i^{t+1} = A_i^t - \frac{2}{3} \Delta A_i^t \ , \tag{12.26}$$

$$A_{i\pm1}^{t+1} = A_{i\pm1}^t + \frac{1}{3} \Delta A_i^t \ . \tag{12.27}$$

Consider now a situation where three neighbouring nodes $(i-1, i, i+1)$ are avalanching; at the central node i A_i will vary according to:

$$A_i^{t+1} = A_i^t + \frac{1}{3} \Delta A_{i-1}^t - \frac{2}{3} \Delta A_i^t + \frac{1}{3} \Delta A_{i+1}^t \ , \tag{12.28}$$

which, upon substitution of eq. (12.25), can be written as

$$\{A\}^{t+1} - \{A\}^t = -\frac{2}{3} [S][S] \{A\}^t \ , \tag{12.29}$$

where $\{A\}^t$ the vector of nodal variables at time t, and $[S]$ a tridiagonal matrix operator defined as

$$S_{ij} = -\frac{1}{2}\delta_{i,j-1} + \delta_{i,j} - \frac{1}{2}\delta_{i,j+1} \ , \tag{12.30}$$

where $\delta_{i,j}$ is the usual Kronecker delta[4]. Now, the matrix operator (12.30) is identical to that arising from second-order centered finite difference discretization of the 1D differential operator $-(1/2)\partial^2/\partial x^2$ on a spatial mesh of constant, unit spacing $\Delta x = 1$; likewise, the LHS of eq. (12.29) can be interpreted as a simple forward-in-time first order discretization of $\partial/\partial t$, for a unit time step $\Delta t = 1$. Consequently, the discrete evolutionary equation (12.28) can be interpreted as a discretization of the following hyperdiffusion equation:

$$\frac{\partial A(x,t)}{\partial t} = D\frac{\partial^4 A(x,t)}{\partial x^4} \ , \tag{12.31}$$

with a hyperdiffusion coefficient $D = 1/6$.

In a more realistic situation where all nodes are not simultaneously avalanching, we have to take into account of the fact this hyperdiffusion coefficient is a function of time and space, being nonzero only at avalanching nodes:

$$D_i^t = \begin{cases} 1/6 & |\Delta A_i| > Z_c \\ 0 & otherwise \end{cases} \tag{12.32}$$

Equation (12.28) must then be replaced by

$$A_i^{t+1} = A_i^t + 2D_{i-1}^t \Delta A_{i-1}^t - 4D_i^t \Delta A_i^t + 2D_{i-1}^t \Delta A_{i+1}^t \ , \tag{12.33}$$

which leads again to a hyperdiffusion equation, this time with a variable diffusion coefficient:

$$\frac{\partial A(x,t)}{\partial t} = \frac{\partial^2}{\partial x^2}\left(D(A)\frac{\partial^2 A(x,t)}{\partial x^2} \right) \ . \tag{12.34}$$

In more than one spatial dimension, and including a spatiotemporally random forcing term $S(\mathbf{x},t)$ acting only when the system is globally stable, one then obtains an evolutionary equation of the general form:

$$\frac{\partial A(\mathbf{x},t)}{\partial t} = \nabla^2 \left(D(A)\nabla^2 A(\mathbf{x},t) \right) + S(\mathbf{x},t) \ . \tag{12.35}$$

now with

$$D(A) = \begin{cases} \nabla^2 A(\mathbf{x},t) & |\Delta A(\mathbf{x})| > Z_c \\ 0 & otherwise \end{cases} \tag{12.36}$$

[4] For a 2D (3D) version of the model, $[S]$ would be a sparse pentadiagonal (heptadiagonal) matrix operator.

Compare this with Lu (1995b)'s model nonlinear diffusion equation (12.22). Bélanger et al. (2007) have solved numerically the 2D form of this generalized hyperdiffusion equation, and demonstrated that the system's behavior in the statistically stationary SOC state, including size distribution of avalanches, is identical to the 2D version of the LH93 sandpile model, even though differences in avalanching behavior are observed in the transient phase towards SOC (compare their Figs. 1 and 2 to Fig. 12.3 herein).

What would then the link between eq. (12.35) and the governing equations of magnetohydrodynamics? There is currently no clean answer to this key question, but Liu et al. (2003) offer the following argument: in the strong MHD turbulence regime expected to characterize the reconnection process, the linear dissipative term in the MHD induction equation is expected to become:

$$\eta \nabla^2 \mathbf{B} \qquad \rightarrow \qquad \eta_{\rm turb} \nabla^2 \mathbf{B} \ , \tag{12.37}$$

where the magnitude of the "turbulent" magnetic diffusivity is set by the vigor of the turbulent flow:

$$\eta_{\rm turb} \sim \langle u^2 \rangle \ , \tag{12.38}$$

where u is the (turbulent) plasma velocity and the angular brackets denote some appropriate averaging operator (Pouquet et al. 1976). Now, eq. (12.37) is second-order in physical space, but in conjunction with eq. (12.38), the magnetic diffusion term as a whole would becomes fourth order in Fourier space, as would the RHS of eq. (12.35). The validity of this conceptual bridge should be investigated through numerical simulations (see also Guo et al. (2012) for a demonstration of hyperdiffusive behavior in magnetic reconnection under different physical regimes).

12.5 Beyond the sandpile

While the physical interpretation and reverse engineering approach to the cellular automata rules of sandpile models may appear convincing, they are, at some level, fundamentally flawed. This is because finite-difference discretization of partial differential equations makes mathematical sense only if the quantities being discretized vary smoothly over the length scale of the spatial mesh used for discretisation. Yet, by the very nature of most SOC sandpile models, including the LH93 model which has been the focus of the preceding section, the thresholded nature of the redistribution mechanism implies that many physical quantities, notably the (hyper)diffusion coefficient, vary discontinuously on the length scale of the lattice. Moreover, the (ad hoc) choice of mapping between the lattice and the numerical mesh, and of finite difference formulae assumed to be represented by the discrete stability and redistribution rules, has a determining influence on the mathematical

form of the reverse-engineered PDEs (compare, e.g., the reverse engineering approaches of Isliker et al. 1998 and Liu et al. 2003)

A distinct approach to the physical interpretation problem has been to design SOC models of solar flares where the dynamical elements are more directly related to the coronal magnetic field and plasma. A few such models are reviewed in this section, after a brief survey of numerical simulations of MHD coronal dynamics directly inspired by Parker's physical scenario.

12.5.1 Numerical simulations

The numerical simulation of solar flares is an extremely active research area and continues to generate a voluminous literature (see Shibata & Magara 2010, and references therein). The foregoing brief overview is restricted to magnetohydrodynamical simulations meant to capture Parker's physical picture of coronal heating by nanoflares, as per Fig. 12.2 herein. Many such simulations have been published, starting with Mikić et al. (1989), Longcope & Sudan (1994), and Galsgaard & Nordlund (1996). The geometrical and physical setup is the same as illustrated in cartoon form on Fig. 12.2: an initially uniform, rectilinear magnetic field threads a cartesian domain, with the normal component of the field held fixed at the bounding planes. Random horizontal displacements are then imposed within these planes, and the simulations track the subsequent evolution of the magnetic field within the computational box by solving the full set of magnetohydrodynamical equations. These pioneering simulations did demonstrate that boundary forcing by random horizontal flows leads to the buildup and dissipation of spatially localized electrical current systems within the domain. However, their unavoidably limited spatial resolution, coupled to the system's natural tendency to build up tangential discontinuities, makes it very difficult even with today's computing hardware to produce very thin current sheets prone to dynamical instabilities, their dissipation taking place instead through simple Ohmic decay, explicit or of numerical origin.

There are a number of ways around the resolution problem, besides waiting for computers to become sufficiently powerful to crack the problem through the brute force approach. Minimally-diffusive numerical algorithms are available in the literature, and in the MHD some have been shown able to track the buildup of current sheets in an accurate and nonlinearly stable manner down to the scale of the spatial mesh (Bhattacharyya et al. 2010). Such algorithms are ideally-suited to simulate Parker's scenario, and should be applied to the problem.

Another approach is to introduce physical simplifications of the governing equations, based on the idea that in a strongly magnetized coronal loop where instability between neighbouring flux strands set in for a relatively small crossing angles (recall Parker's estimate of 14°). The so-called *reduced MHD*

approximation offers a promising avenue towards this goal. The geometrical setup is once again our now familiar unbent coronal loop, with the loop axis oriented in the cartesian z-direction, and the photosphere corresponding to the $z = 0$ and $z = L$ planes, with interior planes perpendicular to the z-axis representing cross-sections of the coronal loop (viz. Fig. 12.2A). In the present context, reduced MHD involves the following assumptions: (1) the magnetic field along the axis of the loop (z-direction) is much stronger than the orthogonal components in the plane perpendicular to the z-axis, denoted by the subscript "$\perp$" in what follows; (2) magnetic field and velocity gradients in the z-direction are much smaller than within the orthogonal planes; (3) kinetic plasma pressure is much smaller than magnetic pressure; and (4) the plasma is incompressible. Expressing the magnetic field in terms of the Alfvén velocity:

$$\mathbf{a} = \frac{\mathbf{B}}{\sqrt{4\pi\rho}} \ , \tag{12.39}$$

the problem reduces to solving the MHD fluid and induction equations for the perpendicular components of the flow and field:

$$\frac{\partial \mathbf{v}_\perp}{\partial t} + (\mathbf{v}_\perp \cdot \nabla_\perp)\mathbf{v}_\perp = a_0 \frac{\partial \mathbf{a}_\perp}{\partial z} + \nu \nabla_\perp^2 \mathbf{v}_\perp + (\mathbf{a}_\perp \cdot \nabla_\perp)\mathbf{a}_\perp - \frac{1}{2}\nabla \mathbf{a}_\perp^2 \ , \tag{12.40}$$

$$\frac{\partial \mathbf{a}_\perp}{\partial t} + (\mathbf{v}_\perp \cdot \nabla_\perp)\mathbf{a}_\perp = a_0 \frac{\partial \mathbf{v}_\perp}{\partial z} + \eta \nabla_\perp^2 \mathbf{a}_\perp + (\mathbf{a}_\perp \cdot \nabla_\perp)\mathbf{v}_\perp \ , \tag{12.41}$$

where a_0 is the strength of the dominant longitudinal (z-directed) magnetic component, and ν and η are the plasma viscosity and magnetic diffusivity, respectively. Two-dimensional numerical solutions of this set of equations subjected to random forcing at the photospheric planes have amply demonstrated that the system reacts by producing numerous, small but intense electrical current sheets within the interior volume, dissipation taking place in a strongly intermittent manner (see, e.g., Dmitruk et al. 1998; Georgoulis et al. 1998; and references therein). However, in these calculations the scale invariance observed in the size distribution of discrete energy release events is a direct reflection of a scale invariance in the size distribution of spatially localised electrical current systems that develop in the simulation, itself a consequence of scale invariance in the plasma turbulence generated by boundary forcing. No *bona fide* avalanches are occurring in these simulations, in the sense that dissipative events do not trigger one another.

If global simulations of the complete Parker scenario remain elusive, numerical simulations have provided much insight on the working of individual elements, in particular at the level of the formation and disruption of electrical current sheets in (relatively) simple sheared magnetic configurations (Dahlburg et al. 2005, 2009; Bhattacharyya et al. 2010). Such simulation have confirmed that sheared magnetic flux system do become unstable beyond a critical shear angle, and have shown that transition to MHD turbulence can

occur through a sequence of at least two distinct instabilities (Dahlburg et al. 2009). The physical underpinning of this key component of Parker's scenario thus appears well established.

12.5.2 SOC in reduced MHD

Yet another approach is to design cellular automata that derive directly from the magnetohydrodynamical equations describing the flaring process. Buchlin et al. (2003; see also Einaudi & Velli 1999) have designed an interesting cellular-automaton-type SOC model inspired by the reduced MHD simulations just discussed. The mathematical structure of eqs. (12.41) is such that the magnetic field can be expressed in terms of a z-directed vector potential through two scalar fields α^s, with $s = \pm$, via the relation

$$\mathbf{v}_\perp + s\mathbf{a}_\perp = \nabla_\perp \times (\alpha^s \hat{z}) \ , \tag{12.42}$$

with the magnetic vector potential and the z-component of the electrical current density given respectively by

$$A_z = (\alpha^+ - \alpha^-)/2 \ , \qquad J_z = -\nabla_\perp^2 A_z \ . \tag{12.43}$$

With the nonlinear and dissipative terms neglected, eqs. (12.41) can be recast in the form of a wave equation for α^s describing the propagation of linear Alfvén waves along the positive and negative z-direction. Any perturbation (i.e., forcing) on the "photospheric" boundaries therefore propagates inwards at a speed $a_0 = B_0/\sqrt{4\pi\rho}$ set by the strength B_0 of the longitudinal magnetic field. The Buchlin et al. (2003) cellular automaton proceeds by propagating boundary forcing inwards by solving the aforementioned wave equation, keeping track at every grid point of the resulting electrical current density. Whenever the latter exceeds some preset threshold, the magnetic field is adjusted so as to bring the current back below the instability threshold, on a timescale δt much shorter than boundary forcing, in the spirit of stop-and-go sandpiles. This is carried out by repeated application of the following evolutionary rule on the vector potential at unstable grid points:

$$A_z(x, y, z, t + \delta t) = A_z(x, y, z, t) - \eta \, \delta t \, J_z(x, y, z, t) \ , \tag{12.44}$$

amounting to a first-order time-explicit, second-order centered-in-space discretization of a linear diffusion equation for the vector potential A_z. The magnetic field and flow are then calculated via eqs. (12.42)–(12.43), and boundary forcing resumes.

This system can be driven to statistically stationary state exhibiting SOC behavior, including scale-invariant energy release events. Comparison of the magnetic field and electrical current structures in the planes orthogonal to the

loop axis, as obtained by the Buchlin et al. (2003) model and in the reduced MHD simulations having inspired the model (Georgoulis et al. 1998), reveal both similarities and differences, notably the (relative) paucity of current-sheet-like structures in the former, as compared to the latter. Nonetheless, the Buchlin et al. (2003) cellular automaton offers a very appealing alternate connection to MHD, which clearly merits further investigation.

12.5.3 Fieldline-based models

Morales & Charbonneau (2008, 2009) have developed an anisotropic cellular automaton model in which the dynamical elements are magnetic flux strands. The geometrical setup and evolutionary rules follows closely Parker's physical picture of stressed coronal loops (Fig. 12.2 herein). A 2D $N \times N$ cartesian lattice is filled with nodes having only top+down connectivity, thus defining a set of N initially vertical lines. Each of these lines is taken to represent a magnetic flux tube, the collection of which making up an unbent coronal loop, with the photosphere corresponding to the top and bottom boundaries. The system is assumed periodic in the horizontal, and is forced by imposing sequential horizontal displacements at randomly selected nodes, the displacement itself having non-zero mean. This process mimics the buildup of stresses and twist within the loop, in response to horizontal photospheric fluid motions. Inevitably, two nodes will end up at the same lattice site, which is taken to correspond to the appearance of a tangential discontinuity/current sheet. When this happens, the angle subtended by the two associated flux strands is calculated, and if this angle exceeds a pre-set threshold value the two flux strands involved are cut-and-spliced, mimicking the topological change mediated by magnetic reconnection, and one node displaced away from the dual occupancy lattice site, in order to restore stability. Figure 12.7 illustrates the working of these operations on a very small lattice, in the early stage of the evolution, when the flux strands have only suffered very little deformation. Avalanching behavior results from the fact that the final nodal displacement step can bring the displaced node to a site already occuppied by a node form another flux strand, as on the example displayed on Figure 12.7; in such a case the cut-and-splice-and-displace stabilization sequence begins anew if the angle at this new crossing site exceeds the stability threshold, and so on until there remain no unstable crossings throughout the lattice as a whole.

The driving mechanism here amounts to a form of one-dimensional biased random walk, with the consequence that the length of the flux strands inexorably increases over time. In the incompressible limit, mass conservation imposes that the product of the strand's cross-section A and length L remains constant. Therefore, after a length increase $L \to L'$, the cross section will be $A' = A(L/L')$. Magnetic flux conservation, in turn, requires $BA = B'A'$. The magnetic energy content of strand k will therefore evolve in time according

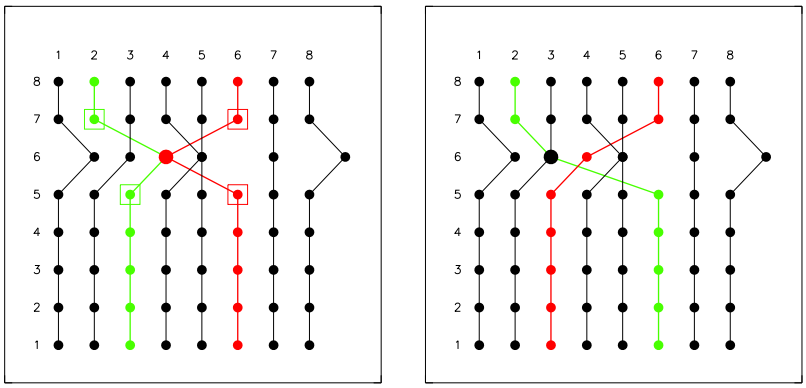


Fig. 12.7: Driving, instability and redistribution mechanism in the anisotropic line-based SOC model of Morales & Charbonneau (2008, 2009). Flux strands are numbered from left to right, and nodes from bottom to top along each strand. On the left panel node (2,6) of strand # 2 (in green) and node (6,6) of strand # 6 (in red) have both been displaced to an empty site initially occupied by a node of strand #4 but left vacant follwing an earlier displacement of node (4,6) to the right. The crossing of strands 2 and 4 defines an angle exceeding here the stability threshold. The right panel shows the result of the cut-and-splice-and-displace redistribution rule, with node (2,6) displaced one lattice unit to the left after the cut-and-splice operation at the unstable lattice site. Should this sequence of operations fail to lower the system's magnetic energy, an additional displacement would be imposed at one of the four nearest neighbours along each flux strand, indicated by square boxes on the left panel (see text). Note also that the displacement of node (2,6) has led to the formation of a new crossing angle between strands # 2 and 3. Adapted from Morales & Charbonneau (2009), Figure 1.

to

$$e_k(t) = \frac{1}{8\pi} \int_{\text{strand}} B^2 \, \mathrm{d}V = \frac{A_0}{8\pi L_0} B_0^2 L_k^2(t) \ , \tag{12.45}$$

where B_0, A_0 and L_0 are the strand's magnetic field strength, cross-section and length in the initial state, and keeping in mind that $\int \mathrm{d}V = A \times L$ here. The total lattice energy is then the sum of the contributions of each strand: $E(t) = \sum e_k(t)$. Clearly, the driving by nodal displacement will increase the system's magnetic energy, and the cut-and-splice-and-displace will tend reduce it[5].

[5] In some instances, the cut-and-splice-and-displace sequence described above does not lead to a reduction of magnetic energy. When this happens, an additional horizontal displacement is imposed at a randomly selected node among the four nearest-

This cellular automaton is readily driven into a SOC state characterized by avalanches of "reconnections" events involving anywhere from 2 flux strands, to the whole lattice. Morales & Charbonneau (2008) demonstrate that the size distribution of energy release events take the form of power laws, with index $\simeq 1.66$ for total liberated energy, $\simeq 1.9$ for peak energy release, and $\simeq 1.9$ for avalanche duration, independently of lattice size and assumed value for the threshold angle. Assuming a loop of 200 G magnetic field strength, length $L_0 = 10^{10}$ cm, and an aspect ratio, 1:100 between the loop's cross-section perimeter to length, the energy released by avalanches spans the range 10^{23}—10^{29} erg for a threshold angle of $\simeq 10°$. Assuming a larger instability threshold angle leads to a lesser energy release, just as in Parker's scenario.

Interpreting the lattice as the outer surface of a coronal loop, Morales and Charbonneau (2009) have synthetized flare "images" by stretching and bending the lattice to yield a more realistic geometry and cross-section to length aspect ratio, followed by projection of the loop so produced on an arbitrarily-oriented observational plane. Synthetic flare images are then constructed by adding the energy released by individual nodes along line-of-sight, in a manner similar to the procedure illustrated on Fig. 12.5. The resulting frequency distribution of flare "areas" as well as their fractal dimension both fall within observational inferences.

Lopez-Fuentes & Klimchuk (2010) have designed an interesting variant on the same idea, in which horizontal displacements are restricted to the bottom boundaries, and fieldlines remain straight as they develop an inclination with respect to the vertical (see their Fig. 2). They also introduce a stability criterion defined in terms of the crossing angle between fieldlines, and introduce discrete redistribution rules that reduce this crossing angle whenever the instability threshold is exceeded. The resulting model is a 1D cellular automaton that releases energy in three more or less distinct stages: growth, statistically steady, and decay. Modelling the plasma response to the associated spatiotemporally intermittent volumetric heating in a simple 1D loop model leads to flare light curves that compare favorably to observations. Although the energy release events do not display the scale-invariance typical of SOC systems, the model illustrates very well the effects of small-scale spatiotemporally intermittent heating on the evolution of coronal loop, in the general context of Parker's physical picture of energy loading and release through forced footpoint motions.

12.5.4 Loop-based models

The SOC model just discussed pertains to tangential discontinuities forming within a single coronal magnetic loop made of many intertwined flux

neighbour of the nodes defining the unstable angle (indicated by square boxes on the left panel of Fig. 12.7). See Morales & Charbonneau (2008) for more detail.

strands, as per Parker's now familiar physical picture. However, high spatial resolution observations reveal that even away from large active region, the solar atmosphere is filled with loops of a wide range of sizes, anchored on photospheric magnetic flux concentrations forming the magnetic network. These loops themselves sometimes cross in response to footpoint displacements forced by photospheric flows, leading to the formation of reconnection-prone tangential discontinuities, now between the surfaces of distinct loops, rather than within them.

Hughes et al. (2003; see also Paczuski & Hughes 2004) have developed a SOC model inspired by this so-called "magnetic carpet" picture of the solar surface (Schrijver et al. 1997; Priest et al. 2002). Dynamical elements are now individual magnetic loops populating a computational photospheric plane. At each temporal iteration, a small random displacement is imposed on a randomly-chosen footpoint (see Fig. 12.8, top panel). With footpoints executing a 2D random walk, the length of the loop connecting them will grow inexorably, leading to an increase of magnetic energy within the loop. Moreover, with many loops present in the system, such random footpoint displacements will inevitably lead to a situation where two loops cross one another, as displayed on the top panel of Fig. 12.8. When this happens, the loops reconnect by exchanging footpoints (middle panel), leading to a reduction of the loop's lengths (bottom panel), which amounts to reducing the magnetic energy within each loop; the excess energy is then released in the corona. Clearly, the reconnection process changes the overall spatial orientation of each loop, and therefore can create new crossing points, leading to new reconnections, and so on until all loop crossings have been eliminated. Reconfigurations can also be triggered by mutual annihilation of two footpoints of opposite polarities, or by merging of footpoints of like polarity. The latter mechanism leads to distinct loops sharing one footpoint. Small loops are removed from the system when their footpoint separation falls below some preset threshold, the model's analog of convective submergence, and the loop population is sustained by injecting small loops at random locations, the analog of flux emergence.

After an initial transient phase, this model settles in a statistically stationary SOC state characterized by an exponential distribution of loop lengths and a power-law distribution for the sizes of energy release events, with logarithmic slope -3. This is steeper than observed, but it remains interesting in that it represents one of the few SOC model of the solar corona that produces a power-law distribution of energy released events steeper than -2, meaning that the smaller events dominate the global heat input into the corona (Hudson 1991), as postulated by Parker in the context of coronal heating (see §12.2 herein).

Figure 12.9 shows a snapshot of the computational plane in the SOC state, and highlights the fact that the population of loops spans a size range going from the smallest loop, barely large enough to avoid elimination as per the submergence criterion, all the way to loops with length comparable to the

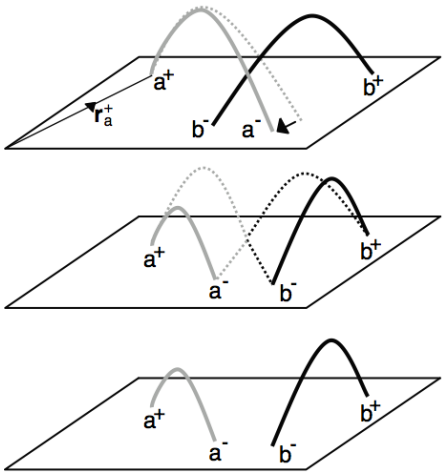


Fig. 12.8: Footpoint displacement and reconnection in the loop-based SOC model of Hughes et al. (2003). Here the random displacement of footpoint a^- (top) has a created a crossing point between the two loops; reconnection (middle) effectively exchanges footpoints a^- and b^-, and leads to energy release through a shortening of the lengths of both loops (bottom). Taken from Hughes et al. (2003), Figure 2.

domain size. The connectivity network between footpoints established by the loop is also scale-free, in the sense that the number of loops tied to each footpoint is also described by a power-law; its logarithmic slope, -2, compares quite well to the distribution inferred by Close et al. (2003) on the basis of force-free reconstructions using photospheric magnetograms.

A very interesting elaboration on this general approach has been subsequently developed by Dimitropoulou et al. (2011). These authors reconstruct nonlinear force-free coronal magnetic fields over temporal sequences of observed magnetograms, and introduce LH93-like cellular automata rules to identify unstable regions and reconfigure this field in response to the evolving lower boundary conditions. A SOC state is eventually reached, characterized by a power-law distribution of energy release events, flatter than in the Hughes et al. (2003) model but in better agreement with the observationally-inferred distribution.

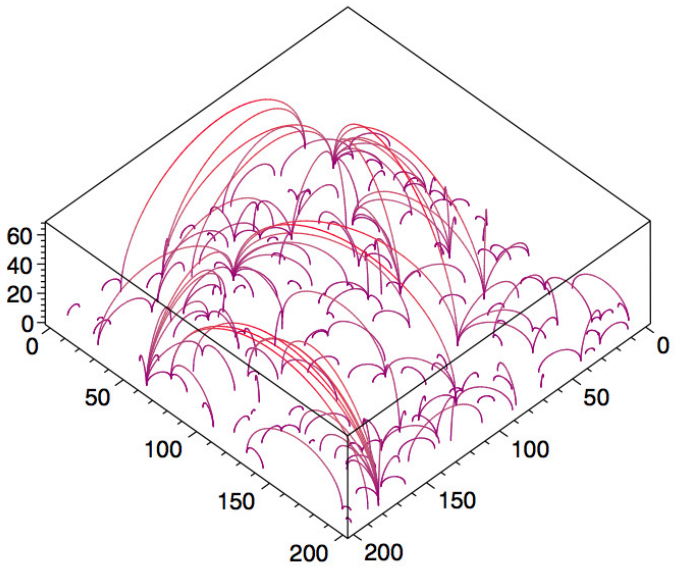


Fig. 12.9: Snapshot of the computational plane in the SOC state, as produced by the model of Hughes et al. (2003). Once in the statistically stationary SOC state, the distribution of loop lengths ranges from the injection scale to the size of the system. Taken from Hughes et al. (2003), Figure 1.

12.6 Outlook

One of the more ambitious "grand challenge" of space weather research is the prediction of large solar flares, with enough lead time to allow mitigation of some of their most dangerous impacts within geospace. What if flares are a manifestation of SOC, along the lines described in this chapter? Under the avalanche *Weltanschau*, the only difference between a large flare and a small one is the number of elementary energy-releasing reconnection events collectively making up the flare; moreover, in both cases the avalanche is triggered by a single current sheet exceeding its stability threshold as the stressed coronal loop adjusts to slow forcing taking place elsewhere in the system, a process taking place at least in part on spatial scales still unaccessible to solar observing platforms. This does not bode well at all for prediction...

It turns out that the situation may not be as dire. If a large avalanche can be triggered in a sandpile, it is because many lattice nodes are very close to their instability threshold; and the distribution of instability thresholds on the lattice is set by the past history of the system, i.e., by avalanche having cascaded through the system at earlier times. Even for a driving process that

is entirely stochastic and on a scale too small to be directly observable, the current state of the lattice does contain information on *potential* avalanches. The numerical experiments of Bélanger et al. (2007) suggest that this information is indeed retrievable through data assimilation of past avalanching behavior, as observed through an energy release time series. In particular, these authors showed that even with purely stochastic spatially uncorrelated driving, production of an optimal initial condition for a 2D LH93-type SOC model by 4DVAR data assimilation can lead to improvements of avalanche forecasts, over direct numerical simulation of the model using an arbitrary initial condition.

Processes or structures exhibiting scale-invariance abound in solar physics, yet in many cases they do not necessarily need be traced to SOC. The outer 30% in radius of the Sun are in a state of strongly turbulent thermally-driven convection, covering a very wide range of spatial scales, including an inertial range spanning many orders of magnitude in wavenumber. That the Sun's magnetic field, imbedded in this strongly turbulent environment, would not be "imprinted" with a corresponding scale invariance would be surprising indeed. Moreover, at photospheric levels processes of advection by convection, aggregation and cancellation can, in themselves, yield scale-invariant distribution of flow and magnetic structures (e.g., Parnell 2001; Rast 2003; Crouch et al. 2007; Thibault et al. 2012). One should certainly not cry "SOC" too swiftly!

The focus of this chapter has been on solar flares, simply because a very convincing physical scenario exists, compatible with the classical requirement for SOC to materialize: a slowly-driven open dissipative system subject to a self-limiting local threshold instability. Moreover, many components of this scenario have received strong support from both observations and numerical simulations. The physical "case" for solar flares as a manifestation of SOC is, arguably, one of the strongest to be found in the geo- and space sciences.

12.7 References

Aschwanden, M.J., Tarbell, T.D., Nightingale, R.W., Schrijver, C.J., Title, A., Kankelborg, C.C., Martens, P., Warren, H.P.: Time variability of the "quiet" sun observed with TRACE. II. Physical parameters, temperature evolution, and energetics of extreme-ultraviolet nanoflares, Astrophys. J. **535**, 1047-1065 (2000)

Aschwanden, M.J.: Physics of the Solar Corona. An Introduction, PRAXIS, Chichester, UK, and Springer, Berlin (2004)

Aschwanden, M.J.: The state of self-organized criticality of the Sun during the last 3 solar cycles. I. Observations, Solar Phys. **274**, 99-117 (2011)

Aschwanden, M.J.: A statistical fractal-diffusive avalanche model of a slowly-driven self-organized criticality system, Astron. Astrophys. **539**, A2 (2012)

Aschwanden, M.J. Freeland, S.L.: Automated solar flare statistics in soft X-rays over 37 years of GOES observations - The Invariance of self-organized criticality during three solar cycles, Astrophys. J. **754**, id112 (2012)

Aschwanden, M.J., Parnell, C.E.: Nanoflare statistics from first principles: fractal geometry and temperature synthesis, Astrophys. J. **572**, 1048-1071 (2002)

Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality - An explanation of 1/f noise, Physical Review Lett. **59/27**, 381-384 (1987)

Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality, Physical Review A. **38/1**, 364-374 (1988)

Bélanger, É., Vincent, A., Charbonneau, P.: Predicting solar flares by data assimilation in avalanche models. I. model design and validation, Solar Phys. **245**, 141-165 (2007)

Benz, A.O.: Flare observations, Liv. Rev. Solar Phys. **5**, lrsp-2008-1 (2008)

Bhattacharyya, R., Low, B,C., Smolarkiewicz, P.K.: On spontaneus formation of current sheets: untwisted magnetic fields, Phys. Plasmas **17**, 112901-112917 (2010)

Buchlin, É., Aletti, V., Galtier, S., Velli, M., Einaudi, G., Vial, J.-C.: A simplified numerical model of coronal energy dissipation based on reduced MHD, Astron. Astrophys. **406**, 1061-1070 (2003)

Charbonneau, P., McIntosh, S.W., Liu, H.L., Bogdan, T.J.: Avalanche models for solar flares, Solar Phys. **203**, 321-353 (2001)

Close, R.M., Parnell, C.E., Mackay, D.H., Priest, E.R.: Statistical flux tube properties of 3D magnetic carpet fields, Solar Phys. **212**, 251-275 (2003)

Crouch, A.D., Charbonneau, P., Thibault, K.: Supergranulation as an emergent length scale, Astrophys. J. **622**, 715-729 (2007)

Dahlburg, R.B., Klimchuk, J.A., Antiochos, S.K.: An explanation for the "switch-on" nature of magnetic energy release and its application to coronal heating, Astrophys. J. **622**, 1191-1201 (2005)

Dahlburg, R.B., Liu, J.-H., Klimchuk, J.A., Nigro, G, S.K.: Explosive instability and coronal heating, Astrophys. J. **704**, 1059-1064 (2009)

Dennis, B.R.: Solar hard X-Ray bursts, Solar Phys. **100**, 465-490 (1985)

Dimitropoulou, M., Isliker, H., Vlahos, L., Georgoulis, M.K.: Simulating flaring events in complex active regions driven by observed magnetograms, Astron. Astrophys. **529**, A101 (2011)

Dmitruk, P., & Gómez, D.O., DeLuca, E.E.: Magnetohydrodynamic turbulence of coronal active regions and the distribution of nanoflares, Astrophys. J. **505**, 974-983 (1998)

Einaudi, G., Velli, M.: The distribution of flares, statistics of magnetohydrodynamic turbulence and coronal heating, Phys. Plasmas, **6**, 4146-4153 (1999)

Galsgaard, K., Nordlund, Å.: Heating and activity of the solar corona. I. Boundary shearing of an initially homogeneous magnetic field, J. Geophys. Res. **101**, 13445-13460 (1996)

Georgoulis, M.K., Velli, M., Einaudi, G.: Statistical properties of magnetic activity in the solar corona, Astrophys. J. **497**, 957-966 (1998)

Guo, Z.B., Diamond, P.A., Wang, X.G.: Magnetic reconnection, helicity dynamics, and hyper-diffusion, Astrophys. J. **757**, 173 (2012)

Hudson, H.S.: Solar flares, microflares, nanoflares, and coronal heating, Solar Phys. **133**, 357-369 (1991)

Hughes, D., Paczuski, M., Dendy, R.O., Helander, P., McClements, K.G.: Solar flares as cascades of reconnecting magnetic loops, Phys. Rev. Lett. **90**(13), id131101 (2003)

Isliker, H., Anastasiadis, A., Vassiliadis, D., Vlahos, L.: Solar flare cellular automata interpreted as discretized MHD equations, Astron. Astrophys. **335**, 1085-1092 (1998)

Isliker, H., Anastasiadis, A., Vlahos, L.: MHD consistent cellular automata (CA) models. I. Basic features, Astron. Astrophys. **363**, 1134-1144 (2000)

Isliker, H., Anastasiadis, A., Vlahos, L.: MHD consistent cellular automata (CA) models. II. Applications to solar flares, Astron. Astrophys. **377**, 1068-1080 (2001)

Jensen, H.J.: Self-Organized Criticality. Emergent complex behavior in physical and biological systems, Cambridge University Press, Cambridge UK, (1998)

Kadanoff, L. P., Nagel, S. R., Wu, L., Zhou, S.: Scaling and universality in avalanches, Phys. Rev. **A39**, 6524-6537 (1989)

Klimchuk, J.A.: On solving the coronal heating problem, Solar Phys. **234**, 41 (2006)

Liu, H., Charbonneau, P., Pouquet, A., Bogdan, T., McIntosh, S.W., Continuum analysis of an avalanche model for solar flares, Phys. Rev. E **66**, id056111 (2002)

Longcope, D.W., Sudan, R.N.: Evolution and statistics of current sheets in coronal magnetic loops, Astrophys. J. **437**, 491-504 (1994)

Lopez-Fuentes, M.C., Klimchuk, J.A.: A simple model for the evolution of multi-stranded coronal loops, Astrophys. J. **719**, 592-601 (2010)

Lu, E.T., Hamilton, R.J.: Avalanches and the distribution of solar flares, Astrophys. J. **380**, L89-L92 (1991)

Lu, E.T., Hamilton, R.J., McTiernan, J.M., Bromund, K.R.: Solar flares and avalanches in driven dissipative systems, Astrophys. J. **412**, 841-852 (1993)

Lu, E.T.: The statistical physics of active regions and the fundamental nature of solar flares, Astrophys. J. Lett. **446**, L109-112 (1995a)

Lu, E.T.: Avalanches in continuum driven dissipative systems, Phys. Rev. Lett. **74**, 2511-2514 (1995b)

Masuda, S., Kosugi, T., Hara, H., Tsuneta, S., Ogawara, Y.: A loop-top hard X-Ray source in a compact solar flare as evidence for magnetic reconnection, Nature **371**, 495-497 (1994)

McIntosh, S.W.: On the inference of differential emission measures using diagnostic line ratios, Astrophys. J. **533**, 1043 (2000)

McIntosh, S. W., Charbonneau, P.: Geometric effects in avalanche models of solar flares: implications for coronal heating, Astrophys. J. Lett. **563**, L165-L168 (2001)

McIntosh, S.W., Charbonneau, P., Bogdan, T.J., Liu, H.-L., Norman, J.P.: Geometric properties of avalanches in self-organized critical models of solar flares, Phys. Rev. E **65**, id046125 (2002)

Mikić, Z., Schnack, D.D., van Hoven, G.: Creation of current filaments in the solar corona, Astrophys. J. **338**, 1148-1157 (1989)

Morales, L. Charbonneau, P.: Self-organized critical model of energy release in an idealized coronal loop, Astrophys. J. **682**, 654-666 (2008)

Morales, L. Charbonneau, P.: Geometrical properties of avalanches in a pseudo-3D coronal loop, Astrophys. J. **698**, 1893-1902 (2009)

Paczuski, M., Hughes, D.: A heavenly example of scale-free networks and self-organized criticality, Physica A **342**, 158-163 (2004)

Parenti, S., Buchlin, E., Cargill, P.J., Galtier, S., Vial, J.-C.: Modelling the radiative signature of turbulent heating in coronal loops, Astron. Astrophys. **651**, 1219-1228 (2006)

Parker, E.N.: Magnetic neutral sheets in evolving fields - Part two - Formation of the solar corona, Astrophys. J. **264**, 642-647 (1983)

Parker, E.N.: Nanoflares and the solar X-Ray corona, Astrophys. J. **330**, 474-479 (1988)

Parnell, C.E.: A model of the solar magnetic carpet, Solar Phys. **200**, 23-45 (2001)

Parnell, C., DeForest, C.E., Hagenaar, H.J., Johnston, B.A., Lamb, D.A., Welsch, B.T.: A power-law distribution of solar magnetic fields over more than five decades in flux, Astrophys. J. **698**, 75-82 (2009)

Pouquet, A., Frisch, U., Leorat, J.: Strong MHD helical turbulence and the nonlinear dynamo effect, J. Fluid Mech. **77**, 321-354 (1976)

Priest, E.R., Heyvaerts, J.F., Title, A.M.: A flux-tube tectonics model for solar coronal heating driven by the magnetic carpet, Astrophys. J. **576**, 533-551 (2002)

Rast, M.P.: The scales of granulation, mesogranulation, and supergranulation, Astrophys. J. **597**, 1200-1210 (2003)

Robinson, P.A.: Scaling properties of self-organized criticality, Phys. Rev. **E49**, 3919-3926 (1994)

Shibata, K., Magara, T.: Solar flares: magnetohydrodynamic processes, Liv. Rev. Solar Phys. **8**, lrsp-2011-6 (2011)

Schrijver, C.J., Title, A.M., van Ballegooijen, A.A., Hagenaar, H.J., Shine, R.A.: Sustaining the quiet photospheric network: the balance of flux emergence, fragmentation, merging and cancellation, Astrophys. J. **487**, 424-436 (1997)

Thibault, K., Charbonneau, P., Crouch, A.D.: The buildup of a scale-free magnetic network, Astrophys. J. **757**, id187 (2012)

Uritsky, V., Paczuski, M., Davila, J.M., Jones, S.I.: Coexistence of self-organized criticality and intermittent turbulence in the solar corona, Phys. Rev. Lett. **99**, id25001 (2007)

Vassiliadis, D., Anastasiadis, A., Georgoulis, M., Vlahos L.: Derivation of solar flare cellular automata models from a subset of the magnetohydrodynamic equations, Astrophys. J. **509**, L53-L56 (1998)

Vlahos, L., Georgoulis, M., Kliuiving, R., Paschos, P.: The statistical flare, Astron. Astrophys., **299**, 897-911 (1995)

Chapter 13
SOC Systems in Astrophysics

Markus J. Aschwanden

The universe is full of nonlinear energy dissipation processes, which occur intermittently, triggered by local instabilities, and can be understood in terms of the *self-organized criticality (SOC)* concept. In Table 2.1 (of chapter 2 of this book) we included a number of cosmic processes with SOC behavior. On the largest scale, galaxy formation may be triggered by gravitational collapses (at least in the top-down scenario), which form concentrations of stars in spiral-like structures due the conservation of the angular momentum. Similarly, stars and planets form randomly by local gravitational collapses of interstellar molecular clouds. Blazars (BL Lac quasars) are active galactic nuclei that have a special geometry with their relativistic jets pointed towards the Earth, producing erratic bursts of synchrotron radiation in radio and X-rays. Soft gamma repeaters are strongly magnetized neutron stars that produce crust quakes (in analogy to earthquakes) caused by magnetic stresses and star crust fractures. Similarly, pulsars emit giant pulses of radio and hard X-ray bursts during time glitches of their otherwise very periodic pulsar signal. Blackhole objects are believed to emit erratic pulses by magnetic instabilities created in the accretion disk due to rotational shear motion. Cosmic ray particles are the result of a long-lasting series of particle acceleration processes accumulated inside and outside of our galaxy, which is manifested in a powerlaw-like energy spectrum extending over more than 10 orders of magnitude. Solar and stellar flares are produced by magnetic reconnection processes, which are observed as impulsive bursts in many wavelengths. Also phenomena in our solar system exhibit powerlaw-like size distributions, such as Saturn ring particles, asteroids, or lunar craters, which are believed to be generated by collisional fragmentation processes or their consequences (in

Markus J. Aschwanden, Solar and Astrophysics Laboratory (LMSAL), Lockheed Martin, Advanced Technology Center, 3251 Hanover St., Palo Alto, CA 94304, USA; e-mail: aschwandenlmsal.com.

Self-Organized Criticality Systems - Dr.Markus J. Aschwanden (Ed.)

form of meteroid impacts). The magnetosphere of planets spawns magnetic reconnection processes also, giving rise to substorms and auroras.

While these astrophysical processes have been interpreted in terms of the self-organized criticality concept (for a comprehensive overview see Aschwanden 2011a), quantitative theoretical modeling of astrophysical SOC phenomena is still largely unexplored. In the following Section 13.1 we outline a general theory approach to SOC phenomena, which consists of a universal (physics-free) mathematical/statistical aspect, as well as a physical aspect that is unique to each astrophysical SOC phenomenon or observed wavelength range. In the subsequent Section 13.2 we discuss then the astrophysical observations and compare the observed size distributions with the theoretically predicted ones.

13.1 Theory

A system with nonlinear energy dissipation governed by self-organized criticality (SOC) is usually modeled by means of cellular automaton (CA) simulations (BTW model; Bak et al. 1987, 1988). A theoretical definition of a SOC system thus can be derived from the mathematical rules of a CA algorithm, which includes: (1) An S-dimensional rectangular lattice grid, (2) a placeholder for a physical quantity $z_{i,j,k}$ associated with each cellular node $x_{i,j,k}$, (3) a definition of a critical threshold z_{crit}, (4) a random input $\Delta z_{i,j,k}$ in space and time; (5) a mathematical re-distribution rule that is applied when a local physical quantity exceeds the critical threshold value which adjusts the state of the nearest-neighbor cells, and (6) iterative time steps to update the system state $z_{i,j,k}(t)$ as a function of time t. Although this definition is sufficient to set up a numerical simulation that mimics the dynamical behavior of a SOC system, it does not quantify the resulting powerlaw-like size distributions in an explicit way, nor does is include any physical scaling law that is involved in the relationship between statistical SOC parameters and astrophysical observables. A quantitative SOC theory should be generalized in such a way that it encompasses both the mathematical/statistical aspects of a SOC system, as well as the physical scaling laws between observables and statistical SOC parameters. In the following we generalize the fractal-diffusive SOC model (FD-SOC), described in Aschwanden (2012a) and outlined in Section 2.2.2, which includes three essential parts: two universal statistical aspects, i.e., (i) the *scale-free probability conjecture*, (ii) the *fractal-diffusive spatio-temporal relationship*, and a physical aspect, i.e., (iii) physical scaling laws between geometric SOC parameters and astrophysical observables, which may be different for each observed SOC phenomenon and each observed wavelength in astrophysical data. Some basic examples of physical scaling laws are derived for fragmentation processes, for thermal emission of astrophysical plasmas, and for astrophysical particle acceleration mechanisms.

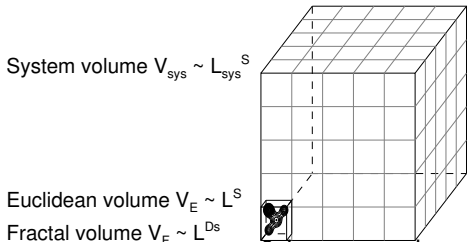


Fig. 13.1: The geometric relationships between the Euclidean avalanche volume $V_E \propto L^S$, the fractal avalanche volume $V \propto L^{D_s}$, and the system volume $V_{sys} \propto L_{sys}^S$ is visualized in 3D space ($S = 3$), as a function of the avalanche length scale L and system size L_{sys}. Note that the probability for an avalanche with size L scales with the ratio of the system volume V_{sys} to the Euclidean avalanche volume V_E.

13.1.1 The Scale-Free Probability Conjecture

Powerlaw-like size distributions are an omnipresent manifestation of SOC phenomena, a property that is also called a "scale-free" parameter distribution, because no preferred scale is singled out by the process. Of course, the scale-free parameter range, over which a size distribution exhibits a powerlaw function, is always limited by instrumental sensitivity or a detection threshold at the lower end, and by the finite length of the time duration over which a SOC system is observed and sampled, at the upper end. Bak et al. (1987, 1988) associated the scale-free behavior with the fundamental property of 1/f-noise that is omnipresent in many physical systems, giving rise to a power spectrum of $P(\nu) \propto \nu^{-1}$.

However, here we give a more elementary explanation for the powerlaw behavior of SOC size distributions, namely in terms of the statistical probability for scale-free avalanche size distributions. A key property of SOC avalanches is that random disturbances can produce both small-scale as well as unpredictable large-scale avalanches of any size, within the limitations of a finite system size L_{sys} at the upper end, and some "atomic" graininess ΔL at the lower end (i.e., a sand grain in sand avalanches, or the spatial pixel size ΔL in a computer lattice grid). The statistical probability distribution $N(L)$ for avalanches with size L can be calculated from the statistical probability. If no particular size is preferred in a scale-free process, the number $N(L)$ of possible avalanches in an S-dimensional system with a volume $V_{sys} = L_{sys}^S$ is simply the system volume V_{sys} divided by the Euclidean volume $V_E = L^S$ of a single avalanche with length scale L (Fig. 13.1),

$$N(L) \propto \left(\frac{L_0}{L}\right)^S \propto L^{-S} \ , \tag{13.1}$$

In a slowly-driven SOC system, none or only one avalanche happens at one particular time, but the relative probability that an avalanche can happen still scales with the reciprocal volume. This is the first basic assumption of our generalized SOC model, which we call the *scale-free probability conjecture*. This conjecture directly predicts a powerlaw function for the basic size distribution of spatial scales, and is distinctly different from the gaussian distribution function that results from binomial probabilities. For 3-dimensional SOC phenomena ($S = 3$), thus we expect a size distribution or differential occurrence frequency distribution of $N(L) \propto L^{-3}$, or a cumulative occurrence frequency distribution of $N(> L) \propto L^{-2}$.

We can now derive the expected size distributions for related geometric parameters, such as for the Euclidean avalanche area A_E or volume V_E. If we simply define the Euclidean area A_E in terms of the squared length scale, i.e., $A_E \propto L^2$ or $L \propto A_E^{1/2}$, which has the derivative $dL/dA_E \propto A_E^{-1/2}$, we obtain for the area size distribution $N(A_E)$, using $N(L) \propto L^{-S}$ (Eq. 13.1),

$$N(A_E)dA_E \propto N[L(A_E)] \left| \frac{dL}{dA_E} \right| dA_E \propto A_E^{-(1+S)/2} dA_E \ , \tag{13.2}$$

yielding $N(A_E) \propto A_E^{-2}$ for 3D phenomena ($S = 3$). Similarly we define the Euclidean volume, i.e., $V_E \propto L^3$ or $L \propto V_E^{1/S}$, which has the derivative $dL/dV_E \propto V^{1/S-1}$, yielding a volume size distribution $N(V_E)$,

$$N(V_E)dV_E \propto N[L(V_E)] \left| \frac{dL}{dV_E} \right| dV_E \propto V_E^{-(2-1/S)} dV_E \ , \tag{13.3}$$

yielding $N(V_E) \propto V_E^{-5/3}$ for 3D phenomena ($S = 3$).

In the case that avalanche volumes are fractal, such as characterized with a Hausdorff dimension D_S in Euclidean space with dimension S, the fractal volume V scales as,

$$V \propto L^{D_S} \ , \tag{13.4}$$

which yields $L \propto V^{1/D_S}$ and the derivative $dL/dV \propto V^{(1/D_S-1)}$, and thus the size distribution,

$$N(V)dV \propto N[L(V)] \left| \frac{dL}{dV} \right| dV \propto V^{-[1+(S-1)/D_S]} \ . \tag{13.5}$$

The Euclidean limit of non-fractal avalanches would yield for $S = 3$ and $D_S = S = 3$ the same powerlaw exponent $\alpha_V = 1 + (S - 1)/D_S = 5/3$ as derived for V_E in Eq. (13.3). For fractal avalanches, with $D_S \approx (1+S)/2 = 2.0$ for $S = 3$, we obtain a slightly steeper powerlaw distribution, $N(V) \propto V^{-2.0}$, than for Euclidean avalanches, i.e., $N(V_E) \propto V_E^{-5/3}$.

A similar effect occurs for fractal avalanche areas A. If we assume an fractal structure with Hausdorff dimension D_2 in 2D space

$$A \propto L^{D_2} \ , \tag{13.6}$$

which yields $L \propto A^{1/D_2}$ and the derivative $dL/dA = A^{1/D_S - 1}$, and thus a size distribution of

$$N(A)dA \propto N[L(A)] \left| \frac{dL}{dA} \right| dA \propto A^{-[1+(S-1)/D_2)]} \ . \tag{13.7}$$

For fractal avalanches in 2D space (S=2), with the fractal dimension $D_2 \approx (1+S)/2 = 1.5$, we obtain a powerlaw distribution $N(A) \propto A^{-5/3}$, which is slightly steeper than for Euclidean avalanche areas, i.e., $N(A_E) \propto A_E^{-3/2}$.

In some astrophysical observations, such as in solar and planetary physics, the areas A of SOC phenomena can be measured, while spatially integrated emission (or spatially unresolved emission) in (optically-thin) soft X-ray or extreme-ultraviolet wavelengths is often roughly proportional to the fractal volume V of the emitting source, and thus the size distributions Eqs. (13.5) and (13.7) can be used to test our scale-free probability conjecture (Eq. 13.1).

If the scale-free probability conjecture (Eq. 13.1) is correct, the derived size distributions for spatial scales $N(L)$, avalanche areas $N(A)$, and avalanche volumes $N(V)$, should be universally valid for SOC phenomena, without any physical scaling laws. They should be equally valid for earthquakes or solar flares, regardless of the physical mechanism that is involved in the non-linear energy dissipation process of a SOC event. In Section 13.2 we will present some astrophysical measurements of size distributions of such geometric parameters (L, A) which can corroborate our assumption of the scale-free probability conjecture. Volume parameters (V) can usually not directly be measured for astrophysical objects, except by means of stereoscopy or tomography of nearby objects.

13.1.2 The Fractal-Diffusive Spatio-Temporal Relationship

After we have established a framework for the statistics of spatial or geometric parameters of SOC avalanche events, we turn now to temporal parameters, which can be defined by a spatio-temporal relationship. The temporal evolution of SOC avalanches is governed by the complexity of nearest-neighbor interactions above some threshold value, which has erratically fluctuating time characteristics according to cellular automaton simulations. However, the mean radius $r(t)$ of an evolving SOC avalanche was found to closely mimic a time dependence of $r(t) \propto t^{1/2}$, which can be associated with a classical random-walk or diffusion process (Aschwanden 2012a). Measurements of the spatial evolution of solar flares, which are considered to be an established

SOC phenomenon, revealed a similar evolution, but tend to be sub-diffusive for the analyzed dataset (Aschwanden 2012b). The observational result of sub-diffusion in solar flares may be related to the highly anisotropic structure of the magnetic field, which inhibits cross-field transport of charged particles, and this way suppresses random walk or classical diffusion, so that the spatially-averaged expansion area corresponds to a sub-diffusive transport process. Moreover, the instantaneous avalanche area was found to have a fractal structure, while the time-integrated avalanche area is nearly space-filling and thus can be described with an Euclidean area or volume. These two properties of fractal geometry and diffusive evolution have been combined in the *fractal-diffusive SOC avalanche model (FD-SOC)* (Aschwanden 2012a). We generalize this concept now also for anomalous diffusion,

$$r(t) \propto \kappa \; t^{\beta/2} \; , \tag{13.8}$$

where κ is the diffusion coefficient and the diffusive exponent β combines classical diffusion ($\beta = 1$), as well as anomalous diffusion ($\beta \neq 1$). Anomalous diffusion processes include both sub-diffusion ($\beta = 0...1$), as well as super-diffusion ($\beta = 1...2$), also called hyper-diffusion or Lévy flights (see also section 2.3.4 in chapter 2 of this book),

$$r(t) \propto t^{\beta/2} \qquad \begin{cases} \beta < 1 \text{ (sub-diffusion)} \\ \beta = 1 \text{ (classical diffusion)} \\ \beta > 1 \text{ (super-diffusion or Lévy flights)} \end{cases} \tag{13.9}$$

We show the generic time evolution of a sub-diffusion process with $\beta = 1/2$ and a super-diffusion process with $\beta = 3/2$ in Fig. 13.2. Anomalous diffusion implies more complex properties of the diffusive medium than a homogeneous structure, which may include an inhomogeneous fluid or fractal properties of the diffusive medium.

The spatio-temporal evolution of an instability generally starts with an exponential growth phase (which we may call the acceleration phase), followed by a saturation or quenching phase (which we may call deceleration phase). In the logistic growth model (Section 2.2.1), the deceleration phase saturates asymptotically at a fixed value, while diffusive models do not converge but slow down progressively with time (Fig. 13.2).

The diffusive scaling (Eq. 13.9) implies then also a statistical correlation between spatial L and temporal scales T,

$$L \propto T^{\beta/2} \; , \tag{13.10}$$

where T is the time duration of a SOC avalanche. From the scale-free probability conjecture $N(L) \propto L^{-S}$ (Eq. 13.1) we can then directly compute the expected occurrence frequency distribution $N(T)$ for time durations T,

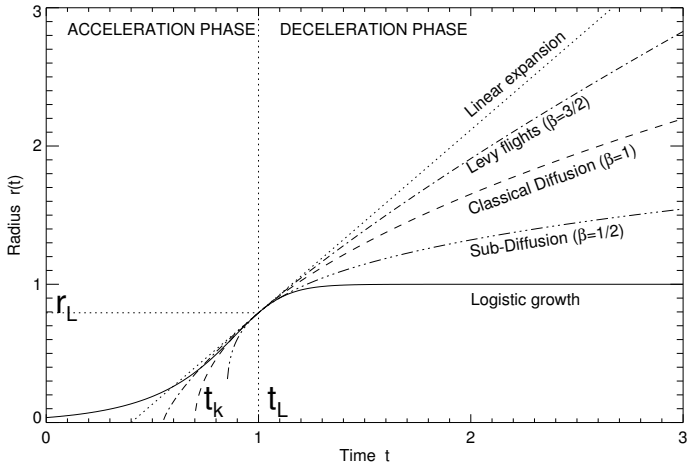


Fig. 13.2: Comparison of spatio-temporal evolution models: Logistic growth with parameters $t_L = 1.0, r_\infty = 1.0, \tau_G = 0.1$, sub-diffusion ($\beta = 1/2$), classical diffusion ($\beta = 1$), Lévy flights or super-diffusion ($\beta = 3/2$), and linear expansion ($r \propto t$). All curves intersect at $t = t_L$ and have the same speed $v = (dr/dt)$ at the intersection point at time $t = t_L$. (Aschwanden 2012b).

$$N(T)dT \propto N[L(T)] \left| \frac{dL}{dT} \right| dT \propto T^{-[1+(S-1)\beta/2]} \, dT \; . \tag{13.11}$$

For instance, for 3D SOC phenomena ($S = 3$) we expect a powerlaw distribution $N(T) \propto T^{-(1+\beta)}$, which amounts to $N(T) \propto T^{-2.0}$ for classical diffusion, $N(T) \propto T^{-1.5}$ for a sub-diffusion case ($\beta = 0.5$), or $N(T) \propto T^{-2.5}$ for a super-diffusion case ($\beta = 1.5$). The exponentially growing phase is neglected in this derivation, which implies a slight underestimate of the number of short time scales. Time scales can also directly be measured in most SOC phenomena, and thus provide an immediat test of the fractal-diffusive assumption made here, regardless of the physical process that is involved in the observed signal of SOC avalanches.

13.1.3 Size Distributions of Astrophysical Observables

The previous theory on geometric (L, A, V) and temporal (T) parameters should be universally valid for the statistics of SOC phenomena, and thus constitutes a purely "physics-free" mathematical or statistical property of SOC systems. All other observables of SOC events, however, are related to a

physical (nonlinear) energy dissipation process, which needs to be modeled in terms of a correlation or scaling law with respect to the physics-free spatio-temporal SOC parameters. Say, if we observe a physical SOC variable x that has a correlation or powerlaw scaling law of $x \propto L^{\gamma}$ with the geometric SOC parameter L, we can infer the expected size distribution $N(x)dx$ by substituting the scaling law.

What is most common in astrophysical observations is the flux F or intensity that is observed in some wavlength range λ (with physical units of energy per time), originating from a source with unknown volume V. The flux $F(t)$ can exhibit strong fluctuations during an energy dissipation event, but we can characterize the time profile of the event with a peak flux P, or with the time-integrated flux, also called fluence (with physical untis of energy), which we may denote with E. For optically-thin emission observed in soft X-ray or EUV emission, the emissivity or flux is approximately proportional to the source volume V, so it is most useful to quantify a scaling law of the flux F with the 3D volume V, which we characterize with a powerlaw exponent γ,

$$F \propto V^{\gamma} \ . \tag{13.12}$$

From the size distribution of the fractal volume $N(V) \propto V^{-[1+(S-1)/D_S]}$ (Eq. 13.5) and the scaling law $V \propto F^{1/\gamma}$ (Eq. 13.12) and its derivative $dV/dF \propto F^{1/\gamma-1}$ we can then derive the size distribution $N(F)$ of fluxes F,

$$N(F)dF \propto N[V(F)] \left| \frac{dV}{dF} \right| dF \propto F^{-[1+(S-1)/(\gamma D_S)]} dF \ . \tag{13.13}$$

which has a typical powerlaw exponent of $\alpha_F \approx 2.0$ (for $S = 3$, $D_S \approx (1+S)/2 \approx 2.0$, and $\gamma \approx 1$). For peak fluxes P we have the same distribution, except for the fractal dimension having its maximum Euclidean value $D_S \approx S$, which yields,

$$N(P)dP \propto N[V(P)] \left| \frac{dV}{dP} \right| dP \propto P^{-[1+(S-1)/\gamma S]} dF \ . \tag{13.14}$$

which has a typical value of $\alpha_P \approx 5/3 \approx 1.67$. Finally, the total flux or fluence $E = \int F(t)dt \approx FT$, is found to have a size distribution of,

$$N(E)dE \propto N[V(E)] \left| \frac{dV}{dE} \right| dE \propto E^{-[1+(S-1)/(\gamma D_S+2/\beta)]} dE \ . \tag{13.15}$$

which has a typical value of $\alpha_E \approx 3/2 = 1.5$, for $S = 3$, $D_S \approx (1+S)/2 \approx 2.0$, $\gamma \approx 1$, and $\beta \approx 1$.

In summary, if we denote the occurrence frequency distributions $N(x)$ of a parameter x with a powerlaw distribution with power exponent α_x,

$$N(x)dx \propto x^{-\alpha_x} \ dx \ , \tag{13.16}$$

we have the following powerlaw coefficients α_x for the parameters $x = L, A, V, T, F, P$, and E,

$$\begin{aligned}
\alpha_L &= S \\
\alpha_A &= 1 + (S-1)/D_2 \\
\alpha_V &= 1 + (S-1)/D_S \\
\alpha_T &= 1 + (S-1)\beta/2 \\
\alpha_F &= 1 + (S-1)/(\gamma D_S) \\
\alpha_P &= 1 + (S-1)/(\gamma S) \\
\alpha_E &= 1 + (S-1)/(\gamma D_S + 2/\beta)
\end{aligned} \quad . \tag{13.17}$$

Thus, the various powerlaw indices depend on four fundamental parameters: the Euclidean dimension S of the SOC system, the fractal dimension D_S of SOC avalanches, the diffusion exponent β of the SOC avalanche evolution, and the scaling law exponent γ between the observed flux and the SOC avalanche volume. Note, that the powerlaw slopes of the geometric (L, A, V) and flux parameters (F, P) do not depend on the diffusion exponent β, and thus are identical for classical or anomalous diffusion. Only the powerlaw slopes of the length scale, time scale, and peak flux (L, T, P) do not depend on the fractal dimension. All flux-related parameters (F, P, E) depend on a physical scaling law (γ), which may be different for every observed wavelength range.

In the following we generally assume 3D SOC phenomena ($S = 3$), for which the fractal dimension can be estimated by the mean value between the minimum and maximum dimension where SOC avalanches can propagate coherently via nearest-neighbor interactions (which limits the minimum fractal dimension to $D_{S,min} \approx 1.0$ and maximum fractal dimension to $D_{max} = S$),

$$D_S \approx \frac{D_{S,min} + D_{s,max}}{2} = \frac{(1+S)}{2} \ , \tag{13.18}$$

which yields $D_3 \approx 2.0$ for $S = 3$ and simplifies the powerlaw indices to

$$\begin{aligned}
\alpha_L &= 2 \\
\alpha_A &= 7/3 \\
\alpha_V &= 2 \\
\alpha_T &= 1 + \beta \approx 2 \\
\alpha_F &= 1 + 1/\gamma \approx 2 \\
\alpha_P &= 1 + 2/(3\gamma) \approx 5/3 \\
\alpha_E &= 1 + 1/(\gamma + 1/\beta) \approx 3/2
\end{aligned} \quad . \tag{13.19}$$

In astrophysics, the distributions of geometric parameters (L, A, V) can only be determined from imaging observations with sufficient spatial resolution (in magnetospheric, heliospheric, and solar physics), while the distributions of all other parameters $(T, F, P,$ and $E)$ can be measured from any non-imaging observations, such as from point-like stellar objects.

13.1.4 Scaling Laws for Thermal Emission of Astrophysical Plasmas

Solar flares and stellar flares are observed in soft X-ray and extreme-ultraviolet (EUV) wavelengths, where the observed intensity is measured in a particular wavelength range λ given by the instrumental filter response function. Soft X-ray and EUV emission is produced by photons via the free-free bremsstrahlung process, free-bound transitions, or radiative recombination. In strong magnetic fields, cyclotron and gyrosynchrotron emission is also produced at radio wavelengths. Soft X-ray and EUV emission occur usually in the optically thin regime, and thus the total emission measure EM, which is proportional to the observed intensity in a given wavelength λ, is proportional to the volume V of the emitting source,

$$F_\lambda \propto EM = \int n_e^2(\mathbf{x})\ dV \propto\ \langle n_e^2 \rangle\ V\ . \tag{13.20}$$

Thus, if the electron density n_e in a source would be constant, or the same among different flare events, we would have just the simple relation $F_\lambda \propto V^\gamma$ (Eq. 13.12) with the scaling law exponent $\gamma = 1$, which is an approximation that is often made. In fact, this is quite a reasonable approximation for measurements with a narrow-band temperature filter, which is sensitive to a particular electron temperature T_e, and thus probes also a particular range of electron densities n_e and plasma pressure $p = 2n_e k_B T_e$ that depend on this temperature T_e, whatever the scaling law between electron temperature T_e and electron density n_e is.

The proportionality constant between the flux intensity F_λ and emission measure EM is dependent on the wavelength range λ, because each wavelength filter is centered around a different temperature range T_e that corresponds to the line formation temperature in the observed wavelengths λ. Physical scaling laws have been derived to quantify the relationship between electron temperature T_e, electron density n_e, and the spatial length scale L_{loop} of coronal loops, e.g., by assuming a balance between the heating rate, conductive, and radiative loss rate (i.e., the so-called RTV law; Rosner, Tucker, and Vaiana 1978), being (in cgs-units),

$$T_e \approx 1400\ (pL_{loop})^{1/3}\ . \tag{13.21}$$

which we can express in terms of the electron density n_e, using the definition of the ideal gas law, $p = 2n_e k_B T_e$,

$$n_e \approx 1.3 \times 10^9 \left(\frac{T_e}{1\ \mathrm{MK}}\right)^2 \left(\frac{L_{loop}}{10^9\ \mathrm{cm}}\right)^{-1} \quad [\mathrm{cm}^{-3}]\ . \tag{13.22}$$

Thus, if there is no particular correlation between the loop length L_{loop} of the densest flare loops (with the highest emission measure) and the volume V of

the active region, the density is only a function of the electron temperature, $n_e \propto T_e^2$. Consequently, if a narrowband temperature filter is used in a soft X-ray or EUV wavelength range, sensitive to a peak temperature T_λ, the corresponding electron density is given in a narrow range also, $n_e \propto T_\lambda^2$, and thus the flux is essentially proportional to the flare volume (according to Eq. 13.20), $F_\lambda \propto V^\gamma$, with a scaling exponent of $\gamma \approx 1$. On the other hand, if a different powerlaw exponent $\gamma \neq 1$ is measured, such an observation would reveal a systematic scaling of some parameters (T_e, n_e, L_{loop}) of the densest flare loops with the size L of the active region.

Another important quantity we want to calculate is the size distribution of thermal energies E_{th}, for which we expect a scaling law of (using Eq. 13.22),

$$E_{th} = 3 n_e k_B T_e V \propto T_e^3 V / L_{loop} \ , \tag{13.23}$$

where the most dominant value of the electron temperature T_e is given by the peak of the *differential emission measure distribution (DEM)*. Observationally, it was found that the DEM peak temperature of a flare scales approximately with the size L of a flare, i.e., $T_e \propto L$ (Aschwanden 1999), which yields with $V \propto L^{D_S}$ (assuming that L_{loop} with the highest emission measure is uncorrelated with the active region size L),

$$E_{th} \propto L^{3+D_S} \ , \tag{13.24}$$

or $E_{th} \propto L^5$ for $D_S \approx 2.0$ and $S = 3$. The size distribution for thermal energies is then expected to be,

$$N(E_{th}) dE_{th} = N(L[E_{th}]) \left| \frac{dL}{dE_{th}} \right| dE_{th} = E_{th}^{-[1+(S-1)/(3+D_S)]} dE_{th} \tag{13.25}$$

which yields $N(E_{th}) \propto E_{th}^{-1.4}$ for $S = 3$ and $D_S \approx 2.0$. However, we have to be aware that this estimate of the thermal energy contained in a flare requires multi-thermal measurements to derive the peak DEM temperature and cannot be obtained from a single narrowband filter measurement. Note that the size distribution of thermal energies with powerlaw exponent $\alpha_{E_{th}} = 1.4$ is very similar to the total energy in photons in any wavelength range, i.e., fluence, $\alpha_E = 1.5$ (Eq. 13.19).

The foregoing model for the thermal energy requires a scaling law between the flare size L and its statistical temperature T_e. In practice, however, thermal energies were often estimated in a limited temperature range from the filter ratio of two narrowband filters. Such filter ratio measurements are sensitive to a particular temperature T_e and electron density n_e (Eq. 13.22), which are then essentially constants in the expression for the thermal energy, and thus the thermal energy is mainly proportional to the volume,

$$E_{th,V} \propto V \ , \tag{13.26}$$

so that the size distribution of thermal energies V is identical to the size distribution of volumes V, which has a powerlaw slope of $\alpha_{th,V} = \alpha_V = 1 + (S-1)/D_3$, yielding values in the range of $\alpha_{th,V} = 1.67 - 2.0$, depending on fractal ($D_3 \approx (1+S)/2 = 2$ for $S = 3$) or Euclidean ($D_3 = S = 3$) volume measurements.

In some studies, the volume is approximated with a "pill-box" geometry, i.e., the product of the measured flare area A with a constant height h along the line-of-sight, $V = Ah$, which makes the volume proportional to the area, and consequently the thermal energy is mainly proportional to the flare area A,

$$E_{th,A} \propto A \ , \tag{13.27}$$

leading to a size distribution of $E_{th,A}$ that is identical with that of the area A, which has the powerlaw slope $\alpha_{th,A} = \alpha_A = 1 + (S-1)/D_2$, yielding values in the range of $\alpha_{th,A} = 2.0 - 2.3$, depending on fractal ($D_2 \approx (1+S)/2 = 1.5$) or Euclidean ($D_2 = 2$) area measurements in $S = 3$ space. Thus, accurate measurements of these size distributions provide a sensitive tool to diagnose the underlying physical scaling laws and model assumptions.

13.1.5 Scaling Laws for Astrophysical Acceleration Mechanisms

Let us consider some simple examples of particle acceleration processes in astrophysical plasmas, such as solar flares, stellar flares, or cosmic rays. The simplest particle acceleration process is a coherent direct current (DC) electric field, which can be characterized by an acceleration constant a over a system length L. Newtonian (non-relativistic) mechanics predicts for an electron with mass m_e a velocity $v = at$ after a distance $L = (1/2)at^2$, which corresponds to a kinetic energy E_L of (where the subscript L refers to the length scale of the accelerator),

$$E_L = \frac{1}{2} m_e v^2 = m_e a L \ , \tag{13.28}$$

which implies a linear energy increase with system length, $E_L \propto L$. Thus, the size distribution of energies E_L is identical with that of length scales L, which we obtain from the scale-free probability conjecture $N(L) \propto L^{-S}$ (Eq. 13.1),

$$N(E_L) dE_L = N(L[E_L]) \left| \frac{dL}{dE_L} \right| dE_L = E_L^{-S} dE_L \ , \tag{13.29}$$

yielding an energy spectrum $N(E_L) \approx E_L^{-3}$ for a 3D Euclidean volume. This simplest case is given by coherent acceleration thoughout the entire source volume.

Alternative particle acceleration mechanisms in astrophysics are stochastic acceleration via wave-particle interactions and shock acceleration. Although the particle orbits in both of these acceleration mechanisms are much more stochastic and random-like, in contrast to the linear path in a DC electric field, the energy gain of the particle may scale with the size of the acceleration region (in which a random walk may occur) and the number of subsequent acceleration regions. For cosmic rays, for instance, which are accelerated over enormous lenghts that make up a substantial fraction of our universe, the energy spectrum has a powerlaw-like slope of $\alpha_L \approx 3$, which is consistent with an energy gain proportional to the combined length of the acceleration path (see section 13.2.11).

Another approach could be made by assuming that the magnetic energy density $dE_B/dV = B^2/8\pi$ per volume element converted into nonthermal particle energy is uncorrelated with the system size L, in which case the total converted magnetic energy just scales with the Euclidean volume $V \propto L^S$ of the energy dissipation process, $E_B \propto V_E \propto L^S$. The corresponding size distribution of magnetic energy E_B is then expected to scale as,

$$N(E_B)dE_B = N(L[E_B)\left|\frac{dL}{dE_B}\right| dE_B = E_B^{-(2-1/S)} dE_B \ , \tag{13.30}$$

which translates into $N(E_B) \propto E_B^{-5/3}$ for a 3D Euclidean volume (S=3). Interestingly, this value is also identical with the size distribution of the peak flux P. Alternative scaling laws using the Alfvénic crossing time through the flare volume have also been considered (e.g, Shibata and Yokoyama 1999, 2002; Nishizuka et al. 2008).

Thus, these acceleration models yield powerlaw size distributions in the range of $\alpha_{E_B} = 1.67$ to $\alpha_{E_L} = 3.0$. The measurement of size distributions of nonthermal energies thus can yield valuable diagnostics about the physical nature of the underlying particle acceleration process.

A summary of all theoretically derived powerlaw indices expected in astrophysical systems is compiled in Table 13.1.

13.2 Observations

We discuss now a number of astrophysical observations with regard to their observed size distributions, which we compare with the foregoing theoretical predictions. A more detailed review of these measurements is given in Aschwanden (2011a; chapters 7 and 8). A list of astrophysical SOC phenomena with their particular sources of free energy or physical mechanisms and instabilities that trigger SOC events is provided in Table 2.1 of chapter 2 of this book (entitled *Theoretical Models of SOC Systems*). Detailed SOC models with explicit physical mechanisms specifying the underlying scaling

Table 13.1: Summary of powerlaw indices predicted in astrophysical systems, as a function of the dimensionality S, the fractal dimension D_S, the diffusion power exponent β, and the energy-volume scaling exponent γ.

Parameter	Powerlaw exponent (general expression)	Powerlaw exponent for $S = 3$, $D_3 = 2$, $D_2 = 3/2$ $\beta = 1$, $\gamma = 1$
Length scale L	$\alpha_L = S$	$\alpha_L = 3$
Area A	$\alpha_A = 1 + (S-1)/D_2$	$\alpha_A = 7/3$
Volume V	$\alpha_V = 1 + (S-1)/D_3$	$\alpha_V = 2$
Time duration T	$\alpha_T = 1 + \beta$	$\alpha_T = 2$
Flux F	$\alpha_F = 1 + 1/\gamma$	$\alpha_F = 2$
Peak flux P	$\alpha_P = 1 + 2/(3\gamma)$	$\alpha_P = 5/3$
Fluence E	$\alpha_E = 1 + 1/(\gamma + 1/\beta)$	$\alpha_E = 3/2$
Emission measure EM_λ	$\alpha_{EM_\lambda} = \alpha_V$	$\alpha_{EM_\lambda} = 2$
Thermal energy E_{th}	$\alpha_{E_{th}} = 1 + (S-1)/(3 + D_S)$	$\alpha_{E_{th}} = 7/5$
Thermal energy $E_{th,A}$	$\alpha_{E_{th,A}} = \alpha_A$	$\alpha_{E_{th,A}} = 7/3$
Thermal energy $E_{th,V}$	$\alpha_{E_{th,V}} = \alpha_V$	$\alpha_{E_{th,V}} = 2$
Linear energy E_L	$\alpha_{E_L} = S$	$\alpha_{E_L} = 3$
Magnetic energy E_B	$\alpha_{E_B} = (2 - 1/S)$	$\alpha_{E_B} = 5/3$

laws between spatial, temporal, and physical parameters are still lacking, but could be worked out for each astrophysical phenomenon as a function of the observed wavelengths using the formal framework outlined in Section 13.1.

13.2.1 Lunar Craters

The size of lunar craters was measured from pictures recorded with the lunar orbiters *Ranger 7, 8, 9* by Cross (1966). A size distribution of 1,600 lunar craters, sampled in the *Mare Tranquillitatis* using data from *Ranger 8*, within a range of 0.56 to 69,000 m, is shown in Fig. 13.3, exhibiting a powerlaw distribution ranging over 5 orders of magnitude with a slope of $\beta \approx 2.0$ for the cumulative occurrence frequency distribution, which translates into a powerlaw slope of $\alpha \approx \beta + 1 \approx 3.0$ for the differential occurrence frequency distribution,

$$N(L) \propto L^{-3} \ . \tag{13.31}$$

This corresponds exactly to our prediction of the scale-free probability conjecture for avalanche events in 3D-space ($S = 3$). A similar powerlaw exponent of $\alpha_L = 2.75$ was also found for the size distribution of meteorites and space debris from man-made rockets and satellites (Fig. 3.11 in Sornette 2004).

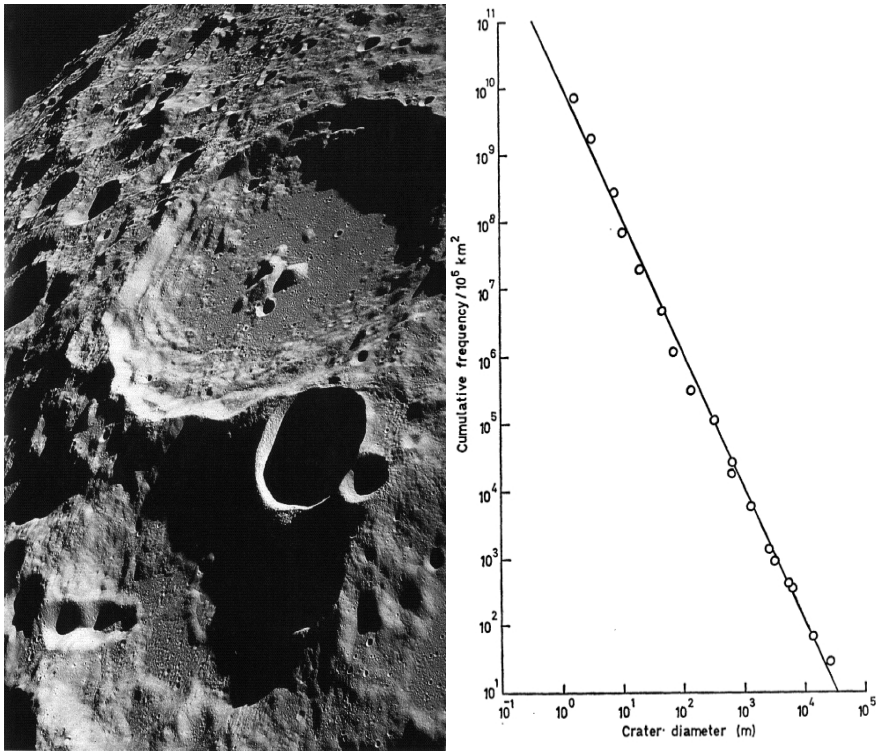


Fig. 13.3: *Left:* The lunar crater Daedalus, about 93 km in diameter, was photographed by the crew of Apollo 11 as they orbited the Moon in 1969 (NASA photo AS11-44-6611). *Right:* Cumulative frequency distribution of crater diameters measured from *Ranger 8* in the lunar *Mare Tranquillitatis* (Cross 1966).

How can we interpret this result? The justification of our scale-free probability conjecture is the fact that the relative probability of partitioning a system into smaller parts scales reciprocally with the volume (Fig. 13.1). The leading theory of lunar crater formation is that their origin was caused by impacts of meteorites, and thus the size distribution reflects that of the impacting meteors and meteorites, which probably were produced by numerous random collisions, similar to the origin of planets, asteroids, and planetesimals. Both the Moon and the Earth were subjected to intense bombardment of solar system bodies between 4.6 and 4.0 billion years ago, which was the final stage of the sweep-up of debris left over from the formation of the solar system. How can we interpret impacting meteorites as a SOC process? The driving force is gravity in the solar system (like gravity drives the infall of sandgrains on a sandpile), while the debris of planetesimals in our solar system represent the sand grains in Bak's sandpile. A self-organizing critical threshold is the result of the combined effect of self-gravity, gravitational

disturbances, collisions, depletions, and captures of incoming new bodies, as discussed in the next section 13.2.2 on the asteroid belt.

An alternative explanation of lunar or terrestrial craters is a volcanic origin. If we attribute volcanoes with nonlinear energy dissipation avalanches in a slowly-driven system of stressing planet crust motions and build-up of subtectonic lava pressure, volcanic eruptions can also be understood as SOC phenomena.

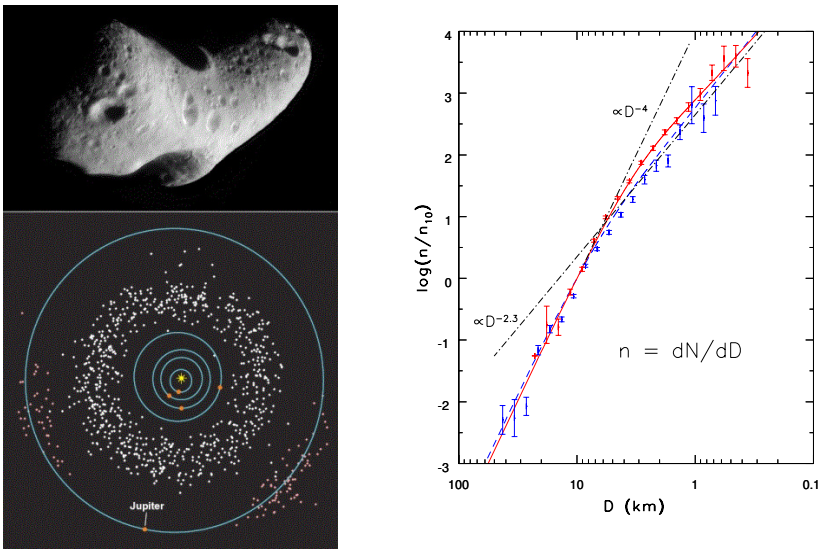


Fig. 13.4: *Left top:* A picture of the near-Earth asteroid *Eros* with a size of 30 km, pictured by a space probe. *Left bottom:* The main asteroid belt located between the Jupiter and Mars orbit. The subgroup of *Trojan asteroids* are leading and trailing along the Jupiter orbit. (Courtesy of NASA/Johns Hopkins University Applied Physics Laboratory). *Right:* Differential size distribution of asteroids observed in the *Sloan Digital Sky Survey* collaboration (Ivezic et al. 2001).

13.2.2 Asteroid Belt

The asteroid size distribution has been studied in the *Palomar Leiden Survey* (Van Houten et al. 1970) and *Spacewatch Surveys* (Jedicke and Metcalfe 1998), where a power law of $N^{cum}(> L) \propto L^{-1.8}$ was found for the cumulative size distribution of larger asteroids ($L > 5$ km), which corresponds to a differential powerlaw slope of $\alpha_L \approx 2.8$. In a *Sloan Digital Sky Survey* collaboration (Fig. 13.4, right), a broken powerlaw was found with $N(L) \propto L^{-2.3}$

for large asteroids (5-50 km) and $N(L) \propto L^{-4}$ for smaller asteroids (0.5-5 km) (Ivezic et al. 2001). In the *Subaru Main-Belt Asteroid Survey*, a cumulative size distribution $N^{cum}(> L) \propto L^{-1.29\pm0.02}$ was found for small asteroids with $L \approx 0.6 - 1.0$ km (Yoshida et al. 2003; Yoshida and Nakamura 2007), which corresponds to a differential powerlaw slope of $\alpha_L \approx 2.3$. Thus, the observed range $\alpha_L \approx 2.3 - 4.0$ of the powerlaw slopes of length scales L is centered around the theoretically expected value of $\alpha_L = 3.0$, predicted by our scale-free probability conjecture.

The origin of asteroids is thought to be a left-over distribution of planetesimals during the formation of planets, which were either too small to form bigger planets by self-gravitation, or orbited in an unstable region of the solar system where Mars and Jupiter constantly cause gravitational disturbances that prevented the formation of another planet. Thus, the final distribution of the asteroid belt is likely to be influenced by both the primordial distribution of the solar system as well as by recent collisions and further fragmentation of planetesimals. The collisional fragmentation process can be considered as a mechanical instability that occurs in a multi-body gravitational field. The collisional process is self-organizing in the sense that the N-body celestial mechanics keeps the structure of the asteroid belt more or less stable, despite of the combined effects of self-gravity, gravitational disturbances, collisions, depletions, and captures of incoming new bodies. The quasi-stability of the asteroid belt warrants the critical threshold in form of a finite collision probability maintained by the proximity of the co-orbiting asteroid bodies.

13.2.3 Saturn Ring

The distribution of particle sizes in Saturn's ring was determined with radio occultation observations using data from the *Voyager 1* spacecraft and a scattering model, which exhibited a powerlaw distribution of $N(r) \propto r^{-3}$ (Fig. 13.5) in the range of 1 mm < r < 20 m (Zebker et al. 1985; French and Nicholson 2000). The size distribution revealed slightly different powerlaw slopes in each ring zone, e.g., $\alpha_L = 2.74 - 3.03$ for ring A, $\alpha_L = 2.79$ for the Cassini division, or $\alpha_L = 3.05 - 3.22$ for ring C (Zebker et al. 1985). These results, again, are consistent with a fragmentation process that obeys the scale-free probability conjecture, similar to the distribution of sizes of asteroids and lunar craters, and predicts a size distribution of $N(L) \propto L^{-3}$. The conclusion that Saturn's ring particles are formed from the (collisional) breakup of larger particles, rather than from original condensation as small particles, was already raised earlier (Greenberg et al. 1977).

The Saturn ring, which mainly consists of particles ranging in size from 1 mm to 10 m, is located at a distance of 7,000-80,000 km above Saturn's equator, and has a total mass of 3×10^{19} kg, just about a little less than the moon Mimas. The origin of the ring is believed to come either from leftover

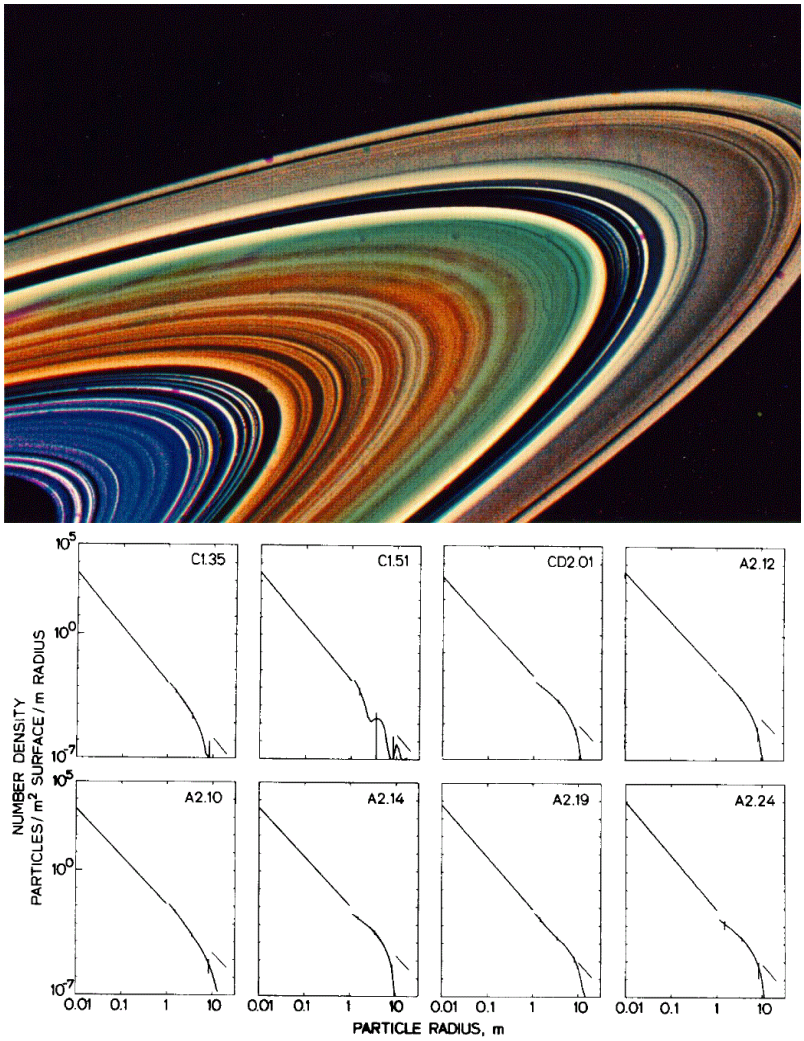


Fig. 13.5: *Top:* Saturn's rings A, B, C, and the Cassini division, photographed by the *Cassini* spacecraft (credit: NASA, JPL, Space Science Institute). *Bottom:* Measurements of the particle size distribution functions for 8 ring regions with *Voyager I* radio occultation measurements (Ring C: C1.35, C1.51; Cassini division: CD2.01; Ring A: A2.12, A2.10, A2.14, A2.19, A2.24). The slopes of the fitted powerlaw functions in these 8 regions are: $\alpha_L = 3.11, 3.05, 2.79, 2.74, 2.70, 2.75, 2.93, 3.03$. The range of particle sizes is $L = 0.01 - 10$ m (Zebker et al. 1985).

material of the formation of Saturn itself, or from the tidal disruptions of a former moon. The celestial mechanics of the Saturn rings is quite complex, revealing numerous gaps in orbits that have harmonic ratios in their periods

with one of the 62 (confirmed) moons (with 13 moons having a size larger than 50 km). Similar to the asteroid belt, the Saturn ring can be considered as a self-organizing system in the sense that the gravity of Saturn and the gravitational disturbances caused by Saturn's moons keep the ring quasi-stable, which provides a critical threshold rate for collisional encounters due to the proximity of the co-orbiting ring particles.

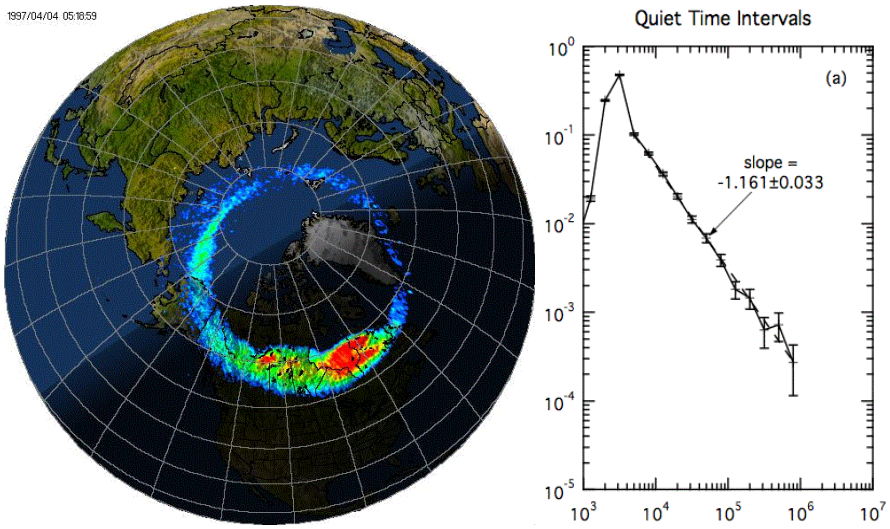


Fig. 13.6: *Left:* Global image of the auroral oval observed by the Ultraviolet Imager (UVI) onboard the NASA satellite "Polar" on April 4, 1997 at 0519 UT, projected onto an Earth map (credit: NASA, Polar/UVI Team, George Parks). *Right:* Occurrence rate frequency distributions of auroral blobs as a function of the area (in units of square kilometers) during substorm-quiet time intervals, recorded with *Polar UVI* during Jan 1-31, 1997 (Lui et al. 2000).

13.2.4 Magnetospheric Substorms and Auroras

The size distribution of auroral areas has been measured with the UV Imager of the *Polar* spacecraft, which exhibits a powerlaw-like distribution with a slope of $\alpha_A = 1.21 \pm 0.08$ during active substorm time intervals, and $\alpha_A = 1.16 \pm 0.03$ during quiescent time intervals (Fig. 13.6; Lui et al. 2000). The corresponding energy flux or power output P of auroral regions was derived to have a powerlaw slope of $\alpha_P = 1.05 \pm 0.08$ during active substorm

time intervals, and $\alpha_P = 1.00 \pm 0.02$ during quiescent time intervals (Lui et al. 2000). These powerlaw slopes are significantly flatter than predicted by our FD-SOC model, i.e., $\alpha_A = 2.33$ and $\alpha_P = 1.67$ for 3D phenomena ($S = 3, D_2 \approx 1.5$). Although Lui et al. (2000) interpret auroras as a SOC phenomenon, the observed powerlaw slopes are far out of the range observed and predicted for other SOC phenomena. Thus, either the sampled distributions are incomplete, they underestimate the areas systematically for smaller events, or SOC models are not applicable for these events.

Plasma flows in the magnetotail plasma with speeds $v \geq 400$ km s^{-1} were found to have a powerlaw distribution of durations T, with $N(T) \propto T^{1.59 \pm 0.07}$ (Angelopoulos et al. 1999), which is not too far off our theoretical prediction (with $\alpha_T = 2.0$), given the relatively small powerlaw range of only ≈ 1.5 decades. Also the size distributions of the durations of AE index ($\alpha = 1.24$; Takalo 1993; Takalo et al. 1999) and AU index ($\alpha = 1.3$; Freeman et al. 2000; Chapman and Watkins 2001) were found to be much flatter than predicted by our SOC model.

Electron bursts in the outer radiation belt (at $4 - 8$ L-shell distances), which may be modulated by fluctuations of the solar wind, were found to have powerlaw distributions with slopes of $\alpha_P = 1.5 - 2.1$ and were interpreted as SOC phenomena (Crosby et al. 2005), which have slopes that are quite consistent with our SOC model ($\alpha_P = 1.67$). The solar wind is thought to be the source of these energetic electrons, although the solar wind velocity frequency distributions were found to exhibit significant deviations from simple powerlaws (Crosby et al. 2005).

13.2.5 Solar Flares

Solar flares are the best studied SOC phenomena in astrophysics. The impulsive energy release associated with solar flares, which can be observed in virtually all wavelengths, from gamma-rays, hard X-rays, soft X-rays, EUV, white-light, infrared, to radio wavelengths, has been interpreted as a SOC phenomen from early on (Lu and Hamilton 1991). Large datasets with $n \approx 10^4 - 10^5$ events sampled over up to eight orders of magnitude in energy provide the necessary statistics to determine accurate slopes of the observed powerlaw-like size distributions. However, major challenges exist still in the elimination of sampling biases in incomplete event sets, the understanding and modeling of powerlaw slopes in different wavelengths in terms of the underlying physical scaling laws, and the automated determination of geometric parameters for large event datasets. A detailed account of observational results sorted into different wavelength regimes is given in Aschwanden (2012a; chapters 7 and 8). We summarize the results of powerlaw slopes observed in size distributions of various SOC parameters in Table 13.2, selecting mostly

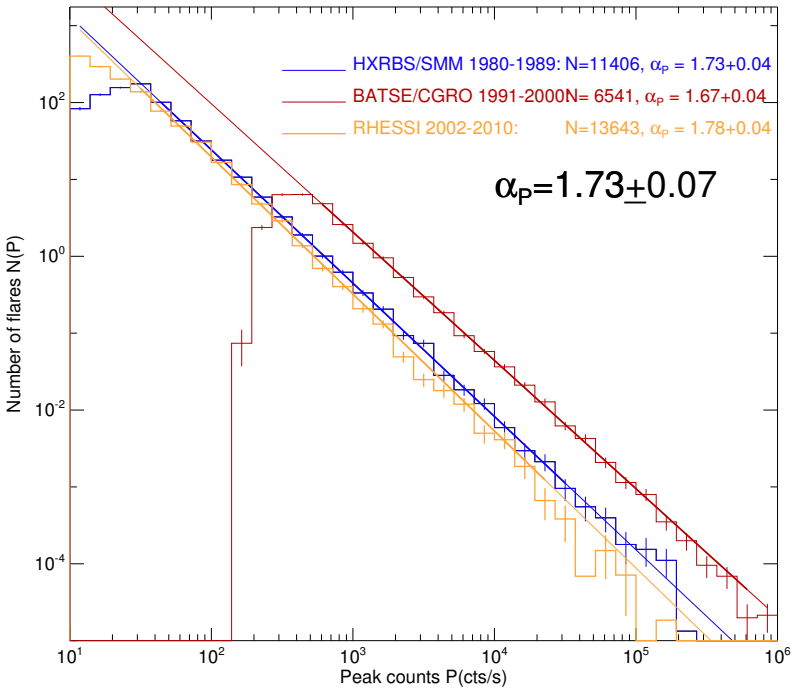


Fig. 13.7: Occurrence frequency distributions of hard X-ray peak count rates P [cts s^{-1}] observed with HXRBS/SMM (1980 – 1989), BATSE (1991 – 2000), and RHESSI (2002 – 2010), with powerlaw fits. An average pre-flare background of 40 [cts s^{-1}] was subtracted from the HXRBS count rates. Note that BATSE/CGRO has larger detector areas, and thus records higher count rates (Aschwanden 2011b).

representative examples with large datasets from different instruments and wavelength regimes.

Size distributions of peak fluxes P (Fig. 13.7), time-integrated fluxes or fluences E (Fig. 13.8), and flare durations T (Fig. 13.9), have been measured for energies > 25 keV in hard X-ray wavelengths with instruments on the spacecraft ISEE-3, SMM, CGRO, and RHESSI, in soft X-ray wavelengths with Yohkoh and GOES, and in EUV with SOHO/EIT, TRACE, and AIA/SDO. Most of the observed powerlaw slopes were measured close to the theoretical predictions, i.e., $\alpha_T = 2.0$, $\alpha_P = 1.67$, and $\alpha_E = 1.5$ (Table 13.2), which is consistent with a dimensionality of $S = 3$, a mean fractal dimension of $D_S \approx (1 + S)/2 = 2.0$, an energy-volume scaling exponent of $\gamma \approx 1$, and a diffusion power exponent of $\beta \approx 1$ (Eq. 13.17 and 13.19).

The measurements in soft X-rays and hard X-rays are all made in broadband energy and wavelength ranges, and thus are least biased regarding a

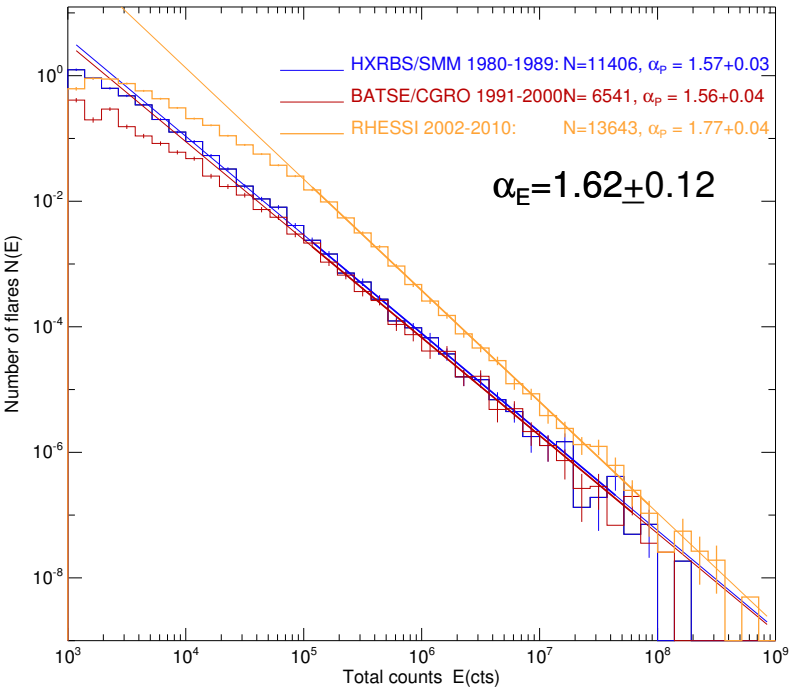


Fig. 13.8: Occurrence frequency distributions of hard X-ray total counts or fluence E [cts] observed with HXRBS/SMM (1980 – 1989), BATSE (1991 – 2000), and RHESSI (2002 – 2010), with powerlaw fits. An average pre-flare background of 40 cts s^{-1} multiplied with the flare duration was subtracted in the total counts of HXRBS (Aschwanden 2011b).

complete sampling of all energy and temperature ranges. The probably most controversial measurements have been made for the smallest flares, also called nanoflares, which have typical temperatures of $T \approx 1-2$ MK and originate in small loops that barely stick out of the transition region. Since EUV measurements with SOHO/EIT, TRACE, and AIA/SDO are all made with narrowband temperature filters, the inferred thermal energies essentially scale with the flare area (Eq. 13.27) or volume (Eq 13.26), for which we predict powerlaw slopes of $\alpha_{th,A} \approx \alpha_A \approx 2.3$ and $\alpha_{th,V} \approx \alpha_V \approx 2.0$, which are significantly steeper than what is predicted for thermal energies sampled with broadband instruments, i.e., $\alpha_{th} \approx 1.5$. This sampling bias has resulted into a controversy whether nanoflares dominate coronal heating, because a powerlaw slope steeper than the critical value of 2 indicates that the energy integral diverges for the smallest events, as pointed out early on (Hudson 1991). Synthesizing measurements from narrowband EUV instruments with broadband soft X-ray instruments, as well as taking the fractal geometry of

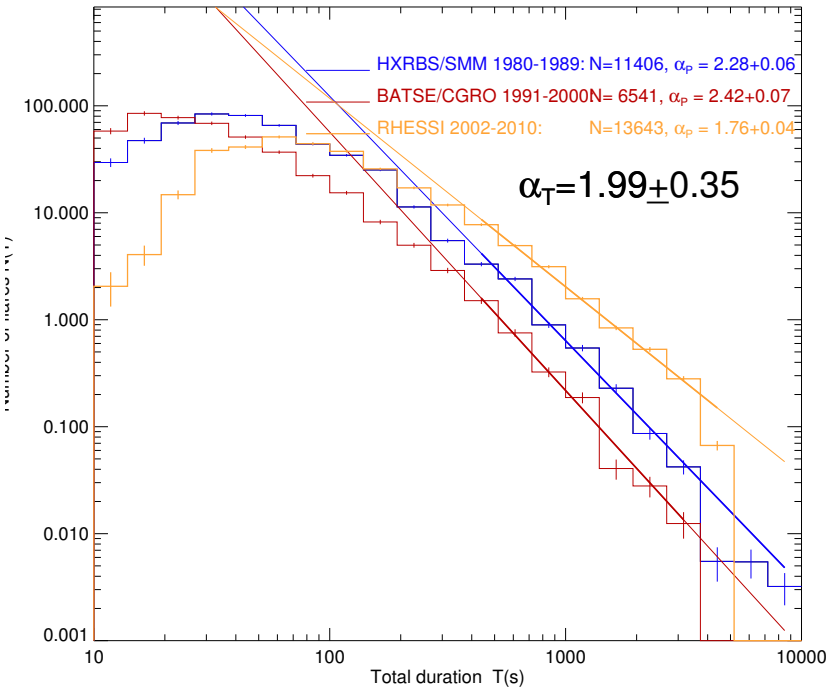


Fig. 13.9: Occurrence frequency distributions of hard X-ray flare durations T [s] observed with HXRBS/SMM (1980 – 1989), BATSE (1991 – 2000), and RHESSI (2002 – 2010) with powerlaw fits. The flare durations for RHESSI were estimated from the time difference between the start and peak time, because RHESSI flare durations were determined at a lower energy of 12 keV (compared with 25 keV for HXRBS and BATSE), where thermal emission prolonges the nonthermal flare duration (Aschwanden 2011b).

flare structures into account, however, could reconcile the size distribution for nanoflares with that of large flares with a corrected value that is close to the theoretical prediction of $\alpha_E \approx 1.5$ (Fig. 13.10; Aschwanden and Parnell 2002).

Problematic are also the measurements of flare durations T for several reasons, such as the limited range of durations over which a powerlaw can be fitted, the ambiguity of separating overlapping long-duration flares, and the solar cycle dependence. While the event overlap problem is not severe during the solar minimum, where a slope close to the theoretically predicted value of $\alpha_T = 2.0$ is measured, the flare event pile-up bias becomes very severe during the solar maximum, producing powerlaw slopes of up to $\alpha_T \lesssim 5$ (Aschwanden and Freeland 2012). Also the powerlaw slope of hard X-ray peak counts α_P

appears to reveal a solar cycle dependence due to a similar effect (Bai 1993, Biesecker 1994, Aschwanden 2011b).

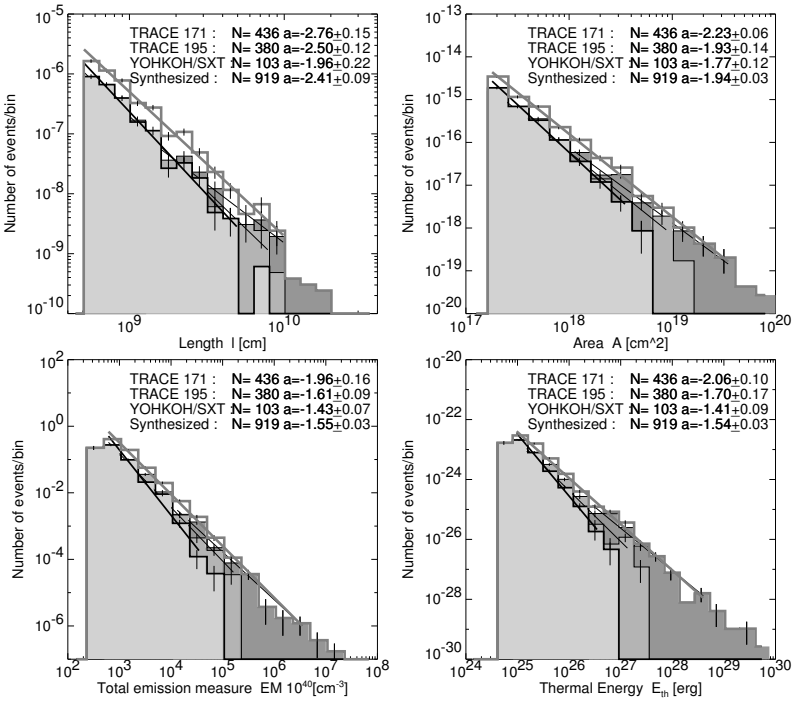


Fig. 13.10: Synthesized frequency distributions from all three wavebands (TRACE 171 Å, 195 Å, and Yohkoh/SXT AlMg) (grey histograms), along with the separate distributions from each waveband (in greyscales). Each of the distributions is fitted with a powerlaw, with the slope values and formal fit errors given in each panel. The four panels represent the four parameters of length L, area A, total emission measure EM (which is proportional to the peak flux P), and the thermal energy E_{th} (Aschwanden and Parnell 2002).

The least explored size distributions of solar flares are the length scale L and area A size distributions. Relatively small samples of flare areas have been measured for EUV nanoflares (Aschwanden et al. 2000, Aschwanden and Parnell 2002) and for the largest M and X-class flares (Aschwanden 2012b). Since the measurement of these parameters provides a direct test of the scale-free probability conjecture (Eq. 13.1), without depending on any other physical parameter or model assumption, priority should be given to such measurements. Existing measurements have large error bars in the powerlaw slope due to the small number of analyzed events, but are largely consistent with the theoretical prediction of $\alpha_L = 3.0$ and $\alpha_A = 2.33$, calculated

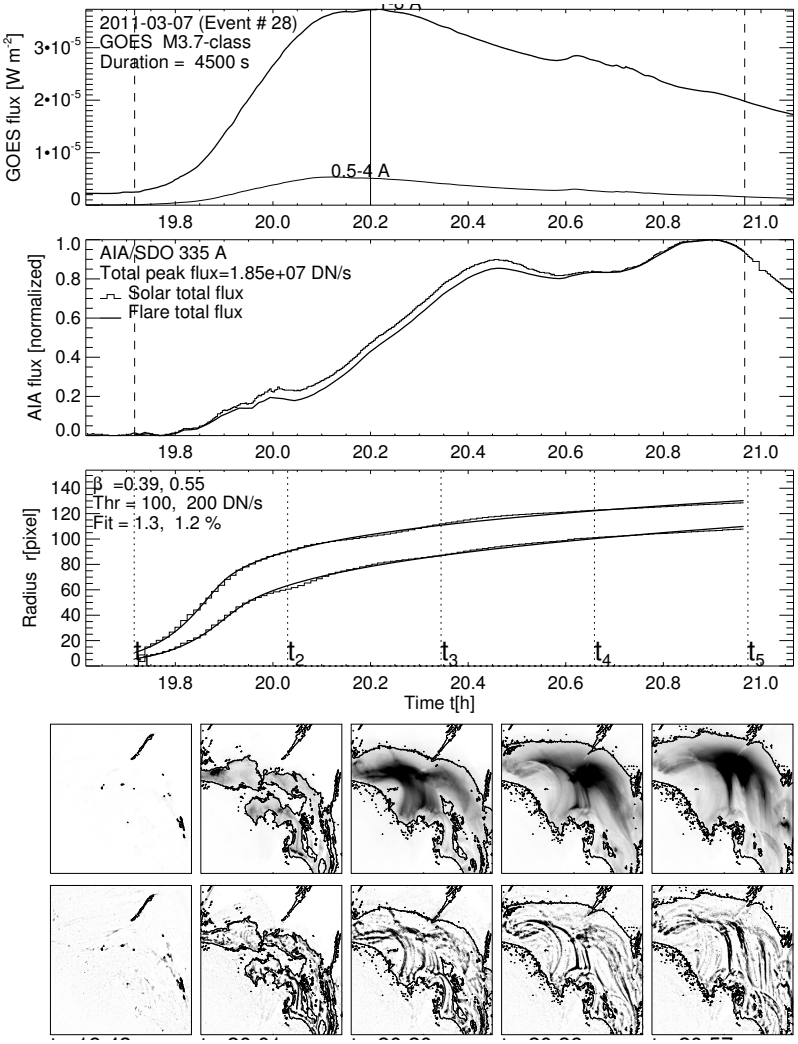


Fig. 13.11: A solar flare event is observed on 2011 Mar 7, 19:43-20:58 UT, with AIA/SDO 335 Å, with GOES time profiles (top panel), the total EUV 335 Å flux (second panel), the spatio-temporal evolution of the radius $r(t) = \sqrt{A(t)/\pi}$ of the time-integrated flare area $A(t)$ for two thresholds, $F_{thresh} = 100, 200$ DN/s (third panel; histogrammed), fitted with the anomalous diffusion model (third panel; solid curve), and 5 snapshots of the baseline-subtracted flux (fourth row) and highpass-filtered flux (bottom row), with the threshold flux $F_{thresh} = 100$ DN/s shown as contour (Aschwanden 2012b).

for 3D SOC avalanches ($S = 3$) with a 2D (area $S = 2$) fractal dimension of $D_2 \approx (1 + S)/2 = 1.5$.

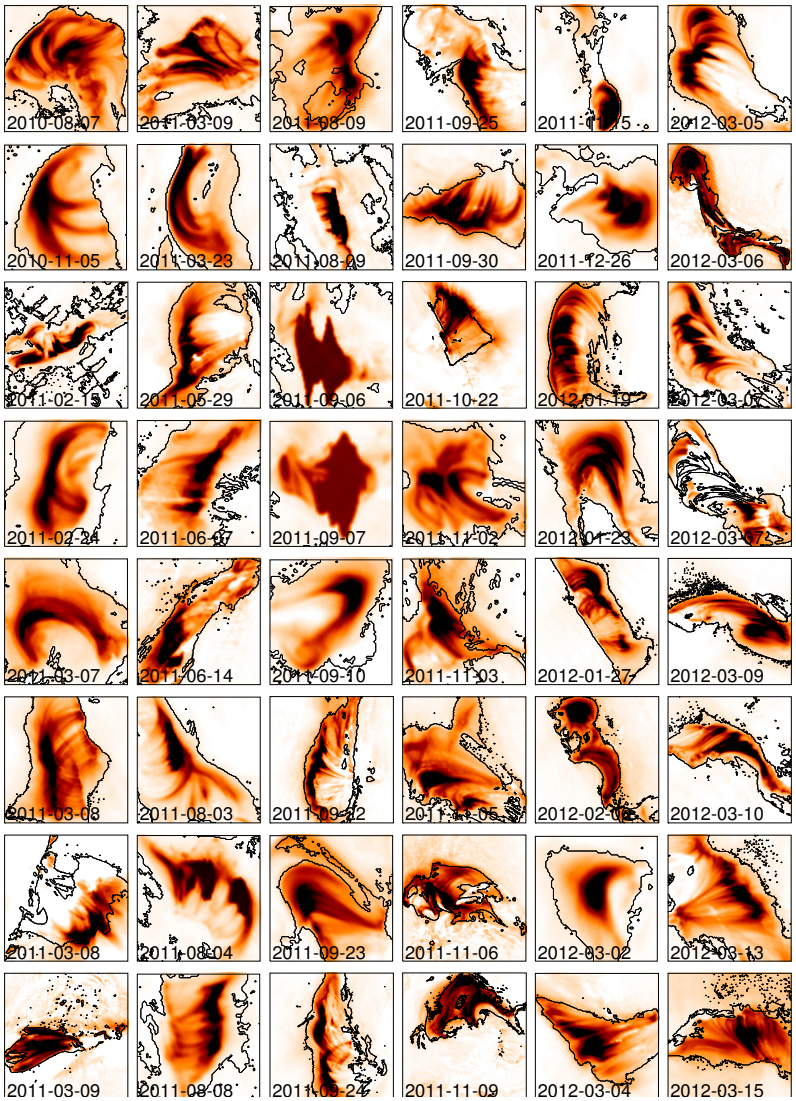


Fig. 13.12: A selection of 48 solar GOES M and X-class flares observed with AIA/SDO at a wavelength of 335 Å during 2010-2012. The contour shows the time-integrated flare area and the color-scale represents the intensity at 335 Å. Note the complex spatial patterns that ressemble to the fractal avalanche patterns of cellular automaton simulations.

Besides the scale-free probability conjecture, a second pillar of our FD-SOC model is the fractal-diffusive spatio-temporal relationship $L(t) \propto \kappa T^{\beta/2}$ (Eq. 13.8), which has been recently tested for a set of the 155 largest (GOES

M and X-class) flare events (Aschwanden 2012b). An example of such a measurement of the spatio-temporal evolution is shown in Fig. 13.11, which exemplifies the spatial complexity of a large cluster of subsequent magnetic reconnection events (Aschwanden 2012b) that form a fractal volume filled with postflare loops (Aschwanden 2012a). Examples of another 48 large flares (GOES M and X-class) are shown in Fig. 13.12. Interestingly, the statistics of these 155 largest flares revealed a sub-diffusive regime ($\beta = 0.53 \pm 0.27$), while classical diffusion appears to be an upper limit. The diffusive characteristics measured in solar flares is consistent with the FD-SOC model, as it is was also found to be consistent with cellular automaton SOC simulations (Aschwanden 2012a).

Another basic ingredient of the fractal-diffusive SOC model is the physical scaling law between the flux observable F_λ (in a given wavelength λ) and the avalanche volume V, which can be expressed by a powerlaw function with coefficient γ, i.e., $F_\lambda \propto V^\gamma$ (Eq. 13.12). This powerlaw exponent γ can only be measured for astrophysical phenomena that are spatially resolved, for instance for solar flares. Statistics of this exponent γ is very sparse in literature, but was found to scale as $F_{335} \propto V^{\gamma_{335}}$ with $\gamma_{335} = 0.79 \pm 0.17$ (Aschwanden 2012b).

Radio bursts are produced in solar flares most frequently either by gyrosynchrotron emission of relativistic electrons that have been accelerated in magnetic reconnection regions (Aschwanden et al. 1993), or by electron beams that escape along magnetic field lines in upward-direction (evidenced by metric type III bursts). Both types of radio bursts (microwave bursts and type III bursts) occur as a consequence of a plasma instability, and thus represent a highly nonlinear energy dissipation process that is typical for SOC processes.

Somewhat out of the predicted range are solar energetic particles, which have rather flat size distributions $N(P)$ of the peak counts, typically in the range of $\alpha_P \approx 1.1 - 1.4$. However, since these are all high-energy particle events (> 10 MeV protons and > 3 MeV electrons), we suspect that only the largest flares produce such high energies, and thus the sample is biased towards the largest flare events, which explains the flatter powerlaw slopes as a consequence of missing weaker events.

13.2.6 Stellar Flares

Impulsive flaring with rapid increases in the brightness in UV or EUV has been observed for a number of so-called flare stars, such as AD Leo, AB Dor, YZ Cmi, EK Dra, or ϵ Eri. These types of stars include cool M dwarfs, brown dwarfs, A-type stars, giants, and binaries in the Hertzsprung-Russell diagram. Most of these stars are believed to have hot soft X-ray emitting coronae, similar to our Sun (a G5 star), and thus magnetic reconnection

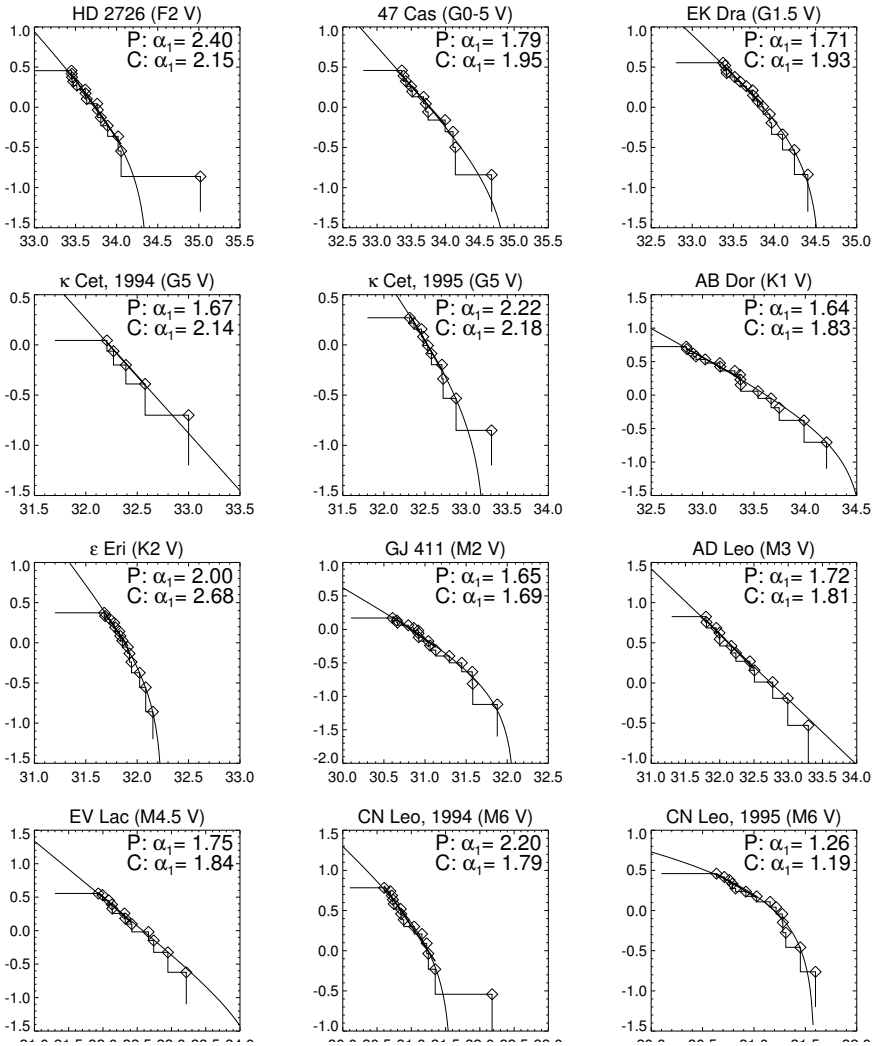


Fig. 13.13: Cumulative frequency distributions of flare energies (total counts) observed for 12 cool (type F to M) stars with EUVE (Audard et al. 2000). The flare events are marked with diamonds, fitted with a powerlaw fit in the lower half (P; thick line), and fitted with a cumulative frequency distribution (C; curved function).

processes are believed to operate in a similar way as on our Sun (see, e.g., review by Guedel 2004).

However, what is different, is that the soft X-ray emission is several orders of magnitude stronger than from our Sun, if we put the stars into the same distance, and thus we expect an observational selection bias towards the

largest possible flares. Another difference in flare statistics is that the meagre observational time allocation in the order of a few hours per star reveals only very few detectable flare events, with typically 5-15 events per observed star (see example in Fig. 13.13). A consequence of this small-number statistics is that we cannot determine a powerlaw slope from a log(N)-log(S) histogram, as we do for larger statistics (of at least 10^2 up to 10^5 events in solar flare data sets), but need to resort to rank-order plots, which correspond to the inverse distribution of cumulative occurrence frequency distributions. So, in principle, an inverse rank-order diagram can be plotted as shown in Fig. 13.13, which shows the logarithmic rank versus the flare energy (or total counts) for each star, from which the powerlaw slope can be determined. However, if we deal with cumulative size distributions, we have also to be aware of the drop-off that results at the upper end of the distribution due to the missing part in the powerlaw differential occurrence frequency distribution above the largest event. Thus, while a straight powerlaw function with slope α can be fitted to a differential frequency distribution $N(E)$,

$$N(E) \propto E^{-\alpha} \ , \tag{13.32}$$

the following function need to be fitted to the cumulative distribution $N_{cum}(> E)$ (Aschwanden 2011a, section 7),

$$N_{cum}(>E) = n\frac{\int_E^{E_{max}} N(E')dE'}{\int_{E_{min}}^{E_{max}} N(E')dE'} = n\frac{\int_E^{E_{max}} E'^{-\alpha}dE'}{\int_{E_{min}}^{E_{max}} E'^{-\alpha}dE'} = n\frac{(E^{1-\alpha} - E_{max}^{1-\alpha})}{(E_{min}^{1-\alpha} - E_{max}^{1-\alpha})} \ . \tag{13.33}$$

The fit of this function to the cumulative distribution is shown in Fig. 13.13 for a set of 12 flare stars, and the resulting values for the powerlaw slope α_E of the inferred differential occurrence frequency distributions (labeled with the letter C in Fig. 13.13). For comparison, we fit also a straight powerlaw with slope $\beta = \alpha - 1$ to the lower half of the cumulative distribution, which is less affected by the upper cutoff (labeled with the letter P in Fig. 13.13). In Table 13.2 we summarize the means and standard deviations of the powerlaw slopes of flare energies observed on flare stars (see individual values in Table 7.7 of Aschwanden 2011a), which have been reported based on various other methods used by the authors, $\alpha_E = 2.17 \pm 0.25$. Fitting only the lower half of the distribution functions we find a significantly lower value of $\alpha_E = 1.85 \pm 0.31$, or by fitting Eq. (13.33) we find a similar range of $\alpha_E = 1.93 \pm 0.35$. This subtle difference in the determination of the powerlaw slope is essential, because it discriminates whether the total energy radiated during stellar flares is dominated by the largest flares (if $\alpha_E < 2.0$) or by nanoflares (if $\alpha_E > 2.0$; Hudson 1991). At this point it is not clear whether the difference in the powerlaw slopes obtained for stellar versus solar flares is due to a methodical problem of small-number statistics, or due to a sampling bias for super-large stellar flares, by solar standards.

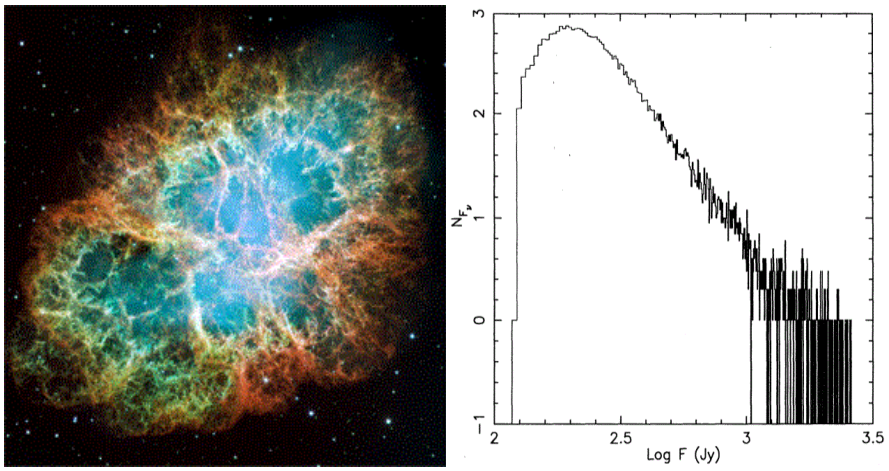


Fig. 13.14: *Left:* Crab nebula, which harbors the pulsar in the center (photographed by Hubble Space Telescope, NASA). *Right:* Frequency distribution of giant-pulse flux densities measured from the Crab pulsar, observed during 15-27 May 1991 with the Green Bank 43-m telescope at 1330, 800, and 812.5 MHz. The tail can be represented by a powerlaw distribution $N_F \propto F^{-\alpha}$ with $\alpha = 3.46 \pm 0.04$ for fluxes $F > 200$ Jy (Lundgren et al. 1995).

13.2.7 Pulsars

Pulsars are fast-spinning neutron stars, which emit strictly periodic signals in radio wavelengths, as well as occasional giant pulses that represent glitches in the otherwise regular pulse amplitude and frequency. The glitches in pulse amplitude and frequency shifts correspond to large positive spin-ups of the neutron star, probably caused by sporadic unpinning of vortices that transfer momentum to the crust. Conservation of the angular momentum produces then an increase of the angular rotation rate. Thus, these giant pulses reveal highly nonlinear energy dissipation processes that can be considered as a SOC phenomenon and we expect a powerlaw function for their size distribution. However, the pulse fluxes are believed to be determined by the magnetospheric physics of the pulsar, while the pulsar glitches are caused by the internal physics of the pulsar crust, which constitutes two different, but possibly linked, SOC systems. Recent SOC models applied to pulsar glitches use a cellular automaton model with superfluid vortex unpinning (Warszawski and Melatos 2008, Melatos and Warszawski 2009, Warszawski and Melatos 2012).

Early measurements of the pulse height distribution of the Crab pulsar (NGC 0532 or PSR B0531+21) observed at 146 MHz were indeed found to have a powerlaw slope of $\alpha_P \approx \beta + 1 = 3.5$ over a range of 2.25 to 300 times

the average pulse size, in a sample of 440 giant pulses (Argyle and Gower 1972). Similar values were measured by Lundgren et al. (1995) in a sample of 30,000 giant pulses, with $\alpha_P \approx 3.06 - 3.36$ (Fig. 13.14, right). While the Crab pulsar is the youngest known pulsar (born in the year 1054), PSR B1937+21 is an older pulsar with a 20 times faster period (1.56 ms) than the Crab pulsar (33 ms). Cognard et al. (1996) measured a powerlaw distribution with a slope of $\alpha_P \approx \beta + 1 = 2.8 \pm 0.1$ from a sample of 60 giant pulses for this pulsar. Statistics of nine other pulsars revealed powerlaw slopes in a much larger range of $\alpha_E = -0.13, ..., 2.4$ for the size distribution of pulse glitches (Melatos et al. 2008), but those measurements were obtained from much smaller samples of 6-30 giant pulses, and thus represent small-number statistics.

The typical value of $\alpha_P \approx 3.0$ found in two pulsars deviates significantly from the prediction ($\alpha_P = 1.67$) of our FD-SOC model, and thus requires either a different model or more statistics from other pulsar cases. Preliminary values of small samples from other pulsars indicate a wide range ($\alpha_E = -0.13, ..., 2.4$, Melatos et al. 2008) that do not point towards a particular value that can be explained with a single model.

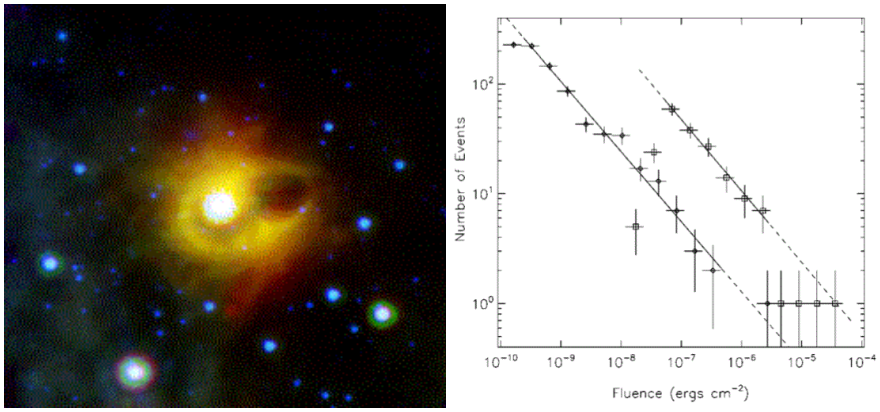


Fig. 13.15: *Left:* The Soft Gamma-Ray Repeater SGR 1900+14 in the constellation of Aquila photographed with the Spitzer Space Telescope in infrared (courtesy: NASA). *Right:* Differential frequency distributions of the fluences of soft gamma-ray repeater SGR 1900+14 observed with CGRO, RXTE, and ICE (Gogus et al. 1999).

13.2.8 Soft Gamma-Ray Repeaters

Gamma-ray bursts were observed from a variety of astrophysical objects, such as neutron stars or black holes, but usually only one burst has been observed from each object. An exception is a class of objects that show repetitive emission at low-energy gamma-rays (> 25 keV), which were termed *soft gamma-ray repeaters (GRS)*. Observations with the *Compton Gamma Ray Observatory (CGRO)* revealed four such SGR sources up to 1999, three in our galaxy and one in the Magellanic Cloud). At least three of these SGR objects were associated with slowly rotating, extremely magnetized neutron stars that are located in supernova remnants (Kouveliotou et al. 1998, 1999). It is believed that these soft gamma-ray bursts occur from neutron star crust fractures driven by the stress of an evolving, ultrastrong magnetic field in the order of $B \gtrsim 10^{14}$ G.

Occurrence frequency distributions of the fluence of soft gamma-ray repeaters were obtained from four SGR sources: a database of 837 gamma-ray bursts from SGR 1900+14 during the 1998-1999 active phase showed a powerlaw slope of $\alpha_E = 1.66$ over 4 orders of magnitude (Fig. 13.15; Gogus et al. 1999); and a combined database from SGR 1806-20, using 290 events detected with the *Rossi X-Ray Timing Explorer*, 111 events detected with *CGRO/BATSE*, and 134 events detected with the *International Cometary Explorer (ICE)*, showing power laws with slopes of α_E=1.43, 1.76, and 1.67 (Gogus et al. 2000). These measurements agree remarkably well with the frequency distributions predicted by the FD-SOC model ($\alpha_E = 1.50$) as well as with those observed during solar flares, which were also observed at the same hard X-ray energies of > 25 keV. However, the physical energy dissipation mechanism may be quite different in a solar-like star and a highly magnetized neutron star, given the huge difference in magnetic field strengths ($B \approx 10^2$ G for solar flares versus $B \approx 10^{14}$ G in a magnetar), although magnetic reconnection processes could be involved in both cases. Nevertheless, soft gamma-ray repeaters have been interpreted as a SOC system (Gogus et al. 1999), in terms of a neutron star crustquake model (Thompson and Duncan 1996), in analogy to the SOC interpretation of earthquakes.

13.2.9 Black Hole Objects

Cygnus X-1, a galactic X-ray source in the constellation Cygnus, is the first X-ray source that has been widely accepted to be a black-hole candidate. The mass of Cygnus X-1 is estimated to be about 14.8 solar masses and it has been inferred that the object with (an event horizon at) a radius of 26 km is far too compact to be a normal star. Cygnus X-1 is a high-mass X-ray binary star system, which draws mass from a blue supergiant variable star (HDE 226868) in an orbit of 0.2 AU around the black hole. The stellar wind

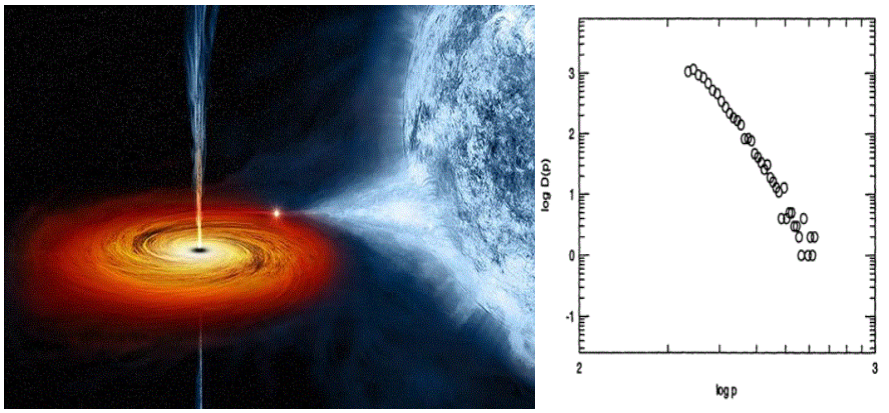


Fig. 13.16: *Left:* This artist concept of Cygnus X-1 shows the black hole drawing material from a companion star (right) into a hot, swirling accretion disk that surrounds the (invisible) black hole (Courtesy of Chandra X-Ray Observatory, NASA). *Right:* Observed frequency distribution of the peak intensities of pulses in the light curve of the black-hole object Cygnus X-1, exhibiting a powerlaw slope of $\alpha_P \approx 7.1$ (Negoro et al. 1995; Mineshige and Negoro 1999).

of this blue companion star swirls mass onto an accretion disk around the black hole (Fig. 13.16). The X-ray time profile from Cygnus X-1 reveals time variability down to 1 ms, which is attributed to X-ray pulses from matter infalling toward the black hole and the resulting turbulence in the accretion disk.

Observations of the X-ray light curve of Cygnus X-1 with Ginga exhibit a complex power spectrum that entails at least 3 piece-wise powerlaw sections, which have been interpreted as a superposition of multiple 1/f-noise spectra (Takeuchi et al. 1995). The occurrence frequency distribution of peak intensities shows a powerlaw-like function with a steep slope of $\alpha_P \approx 7.1$ (Fig. 13.16 right; Negoro et al. 1995; Mineshige and Negoro 1999). All these properties have been modeled with a sophisticated cellular automaton model in the framework of the SOC concept. Infalling mass lumps in the accretion disk are thought to trigger turbulent instabilities in the neighborhood of an infall site, which propagate avalanche-like and produce hard X-rays either by collisional bremsstrahlung or some other magnetically driven instability (e.g., a magnetic reconnection process). The cellular automaton simulations (Takeuchi et al. 1995, Mineshige and Negoro 1999) were able to reproduce steep powerlaw slopes of the peak fluxes in the range of $\alpha_P \approx 5.6 - 11.5$, depending on the effect of enhanced mass transfer by gradual diffusion in addition to the avalanche-like shots, and this way could reproduce the observations.

We note that the observed steep powerlaw slopes of peak fluxes ($\alpha_P \approx 7.1$) exceed the predictions of the FD-SOC model ($\alpha_P = 1.67$) by far. Such a steep slope, $\alpha_P = 1 + 2/(3\gamma) \approx 7$, can be produced by an extremely weak dependence of the X-ray flux F on the avalanche volume V, i.e., $F \propto V^{\gamma}$ with $\gamma \approx 1/9$, which is different from the flux-volume scaling law of optically-thin soft X-ray or EUV emission ($\gamma \approx 1$) generally observed in astrophysical sources, and thus may indicate a nonthermal emission mechanism. Alternatively, the very steep powerlaw slope of α_P could by part of an exponential frequency distribution near the upper cutoff, which could only be proven by sampling peak fluxes with higher sensitivity. Simultaneous modeling of the observed occurrence frequency distributions of time scales, peak fluxes, and power spectra in terms of a SOC model may reveal the underlying physical scaling law of the emission mechanism in black-hole accretion disks.

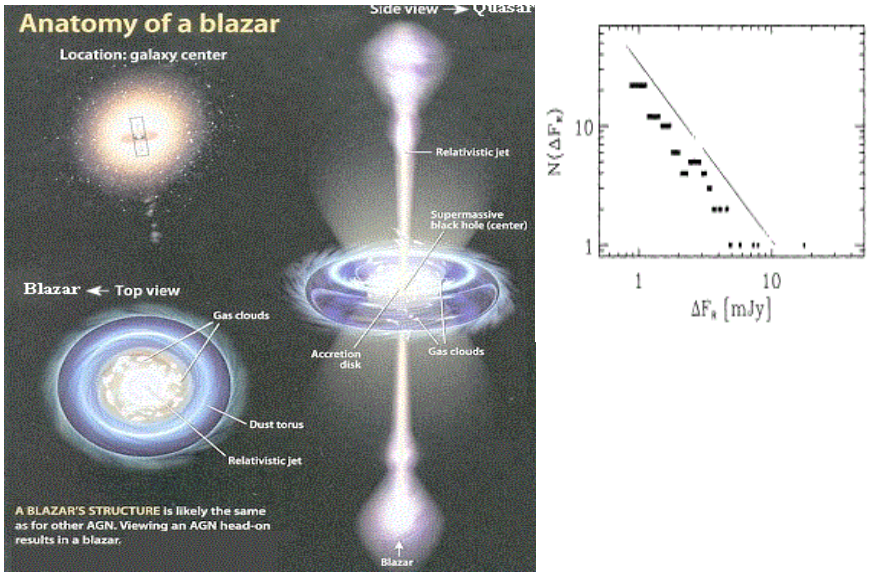


Fig. 13.17: *Left:* The anatomy of a blazar is shown in a top view and side view, having the rotation axis with the relativistic jet oriented toward Earth or orthogonal to the line-of-sight (Courtesy of The Encylopedia of Science). *Right:* Frequency distribution of peak fluxes of flaring events in blazar GC 0109+224, including fluxes above a 3σ-threshold, fitted with a powerlaw function $N(P) \propto P^{-1.55}$ (Ciprini et al. 2003).

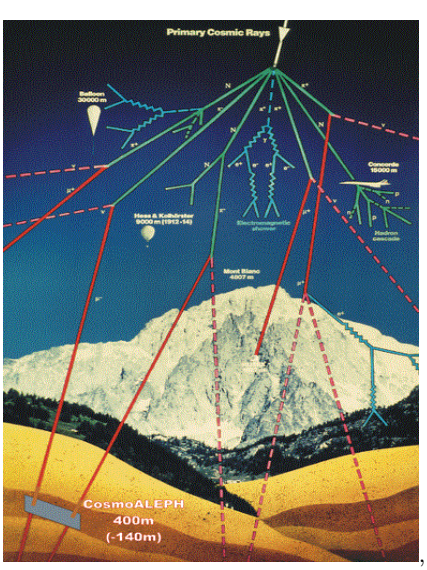
13.2.10 Blazars

Blazars (BL Lac quasars) are a special subgroup of quasars. This group includes BL Lacertae objects, high polarization quasars, and optically violent variables. They are believed to be active galactic nuclei whose jets are aligned within $\lesssim 10^\circ$ of our line-of-sight (Fig. 13.17 left). Because we are observing "down" the jet direction we observe a large degree of variability and apparent superluminal speeds from the jet-aligned emission. The structure of blazars, like all active galactic nuclie, is thought to be powered by material falling onto the supermassive black hole at the center of the host galaxy, which emits highly intermittent gyrosynchrotron emission (in radio wavelengths), inverse Compton emission (in X-rays and gamma-rays), and free-free bremsstrahlung emission (in soft X-rays), modulated by the variable rate of matter infalling into the accretion disk of the central black hole.

The optical variability of blazar GX 0109+224 was monitored and the light curves were found to exhibit flickering and shot noise, with a power spectrum $P(\nu) \propto \nu^{-p}$ with power exponent of $p = 1.57 - 2.05$ (Ciprini et al. 2003). The occurrence frequency distribution of peak fluxes of flare events was found to have a powerlaw slope of $\alpha_p = 1.55$ (Fig. 13.17, right; Ciprini et al. 2003), which is close to the prediction of the FD-SOC model ($\alpha_p = 1.67$). Thus, the highly-variable blazar emission was interpreted in terms of SOC models. The fact that the peak size distribution of radio emission observed in the blazar agrees with the prediction of the FD-SOC model is consistent with a near-proportional radio flux-volume scaling, i.e., $F \propto V^\gamma$ with $\gamma \approx 1$, which is generally the case for gyro-synchrotron emission. This is different from the flux scaling of emission observed from the black hole Cygnus X-1. Thus, SOC statistics allows us to discriminate between different physical emission mechanisms in black holes and blazars.

13.2.11 Cosmic Rays

Cosmic rays are high-energy particles that have been accelerated during a long journey through a large part of our universe, inside our Milky way as well as from outside of our galexy. Their boosting to the highest energies of up to $\lesssim 10^{21}$ eV can only occur by multi-step acceleration processes throughout the universe over long time spans. The energy spectrum of cosmic rays, as it can be measured from showers of secondary particles produced in the Earth atmosphere (Fig. 13.18 left), exhibits a powerlaw-like energy spectrum extending from $\approx 10^9$ eV (=1 GeV) over 12 orders of magnitude up to $\lesssim 10^{21}$ erg (Fig. 13.18 right). The average slope is $\alpha_E \approx 2.7$. However, a more accurate model is a broken powerlaw with a "knee" in the spectrum around $E \approx 10^{16}$ eV. The widely accepted interpretation of this knee is that it separates the origin of cosmic rays from inside and outside of our



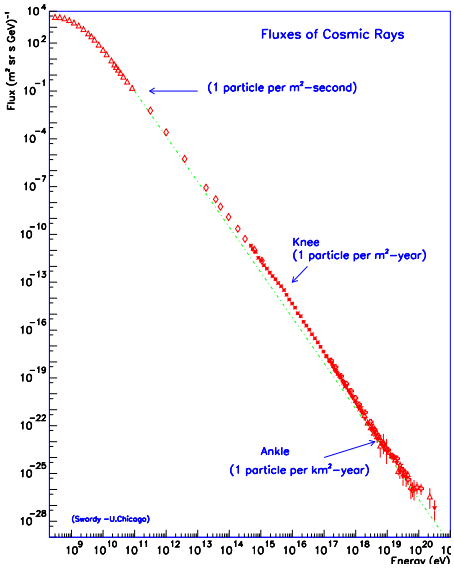


Fig. 13.18: *Left:* The diagram shows the cascade shower of a primary cosmic ray particle, which produces in the Earth's atmosphere a shower of secondary particles, which can be detected by an unterground detector, such as CosmoALEPH (courtesy of CosmoALEPH Team). *Right:* Cosmic ray spectrum in the energy range of $E = 10^9 - 10^{21}$ eV, covering over 12 orders of magnitude. There is a "knee" in the spectrum around $E \approx 10^{16}$ eV, which separates cosmic rays originating within our galaxy (at lower energies) and those from outside the galaxy (at higher energies). (Credit: Simon Swordy, University of Chicago).

galaxy. The powerlaw slope above the "knee" steepens to $\alpha_E \approx 3.0 - 3.3$. Interestingly, our FD-SOC model applied to the kinetic energy gain in a coherent direct current (DC) electric field implies a proportionality of $E_L \propto L$ (Eq. 13.28) for sub-relativistic energies, and thus predicts an energy distribution of $N(E_L) \propto E_L^{-3}$, similar to the cosmic ray spectrum. Of course, cosmic rays are highly relativistic and are likely to be produced by many (n) acceleration phases. However, even if each acceleration phase is local and has a relatively small length of $l \ll L$, the energy gain of the particle would add up linearly with increasing travel time and travel distance $\sum l$ and could still fulfill the proportionality,

$$E_L \propto \sum_{i=1}^{n} E_l \propto \sum_{i=1}^{n} l \propto L \ , \tag{13.34}$$

and end up with an energy spectrum of $N(E_L) \propto L^{-3}$.

So, can we understand the acceleration of a cosmic ray as a SOC process? The lattice grid would cover a large fraction of the extragalactic space of our universe, the ensemble of cosmic-ray particles would represent an avalanche that nonlinearly dissipates energy from local acceleration processes (such as by the first-order Fermi acceleration process between intergalactic or interstellar magnetic clouds), which are self-organizing in the sense that accelerating fields (of magnetic clouds) are constantly restored by the galactic and interstellar dynamics.

Our fractal-diffusive spatio-temporal relationship ($r(t) \propto t^{\beta/2}$, Eq. 13.8) further predicts a random-walk through the universe and an age T that scales as $L \propto T^{1/2}$ with the straight travel distance L from the point of origin of the cosmic ray particle. Using the "knee" in the cosmic-ray energy spectrum (at $E_{gal} \approx 10^{16}$ eV) as a calibration for the distance of the Earth to the center of our galaxy ($L_{gal} \approx 50$ light years $\approx 5 \times 10^{22}$ cm), we can estimate the straight length scale over which the cosmic ray particle travelled by random walk

$$L_{max} \approx L_{gal} \left(\frac{E_{max}}{E_{gal}} \right) \ , \tag{13.35}$$

which yields $L_{max} \approx 5 \times 10^{27}$ erg, which corresponds to about 10% of the size of our universe $r_{uni} \approx 4 \times 10^{28}$ cm. Since the acceleration efficiency is different in galactic and extragalactic space, the diffusion coefficient of the random walk is also different and therefore we expect a different powerlaw slope in the energy spectrum produced in these two regimes.

13.3 Conclusions

In this chapter we generalized the fractal-diffusive self-organized criticality (FD-SOC) model in terms of four fundamental parameters: (i) the Euclidean dimension S, (ii) the fractal dimension D_S of the spatial SOC avalanche structure, (iii) the diffusion exponent β that includes both sub-diffusion and super-diffusion, and (iv) the energy-volume scaling law with powerlaw exponent γ. This model predicts powerlaw functions for the occurrence frequency distributions of the SOC model, and moreover predicts their powerlaw slope as a function of the four fundamental parameters. For a Euclidean dimension of $S = 3$, a mean fractal dimension of $D_S \approx (1 + S)/2 = 2$, classical diffusion ($\beta = 1$), and linear flux-volume scaling ($\gamma = 1$), our generalized FD-SOC model predicts then the following powerlaw slopes: $\alpha_L = 3$ for length scales, $\alpha_A = 2.333$ for areas, $\alpha_T = 2$ for durations, $\alpha = 1.667$ for peak fluxes, and $\alpha_E = 1.5$ for fluences or total energies of the SOC avalanches.

Comparing these theoretical predictions with the observed powerlaws of size distributions in astrophysical systems (summarized in Table 13.2) we find acceptable agreement for the cases of lunar craters, asteroid belts, Saturn rings, outer radiation belt electron bursts, solar flares, soft gamma-ray re-

peaters, and blazars, if we apply the linear flux-volume scaling. Discrepancies are found for magnetospheric substorms, stellar flares, pulsar glitches, black holes, and cosmic rays, which apparently require a nonlinear flux-volume scaling. Pulsar glitches and cosmic rays can indeed be modeled by assuming a linear energy-length scaling, which leads to energy spectra of $N(E) \approx E^{-3}$. Black-hole pulses have very steep size spectra, which indicates a quenching or saturation process that prevents a large variation of pulse amplitudes. Magnetospheric substorms and solar energetic particles have the flattest size distributions, which possibly can be explained by a selection effect with a bias for the largest events. In conclusion, the generalized FD-SOC model can explain a large number of astrophysical observations and can discriminate between different scaling laws of astrophysical observables. We envision that more refined scaling laws between astrophysical observables will be developed that are consistent with the observed size distributions, and this way will provide the ultimate predictive power of SOC models.

Table 13.2: Summary of theoretically predicted and observed powerlaw indices of size distributions in astrophysical systems.

	Length α_L	Area α_A, $\alpha_{th,A}$	Duration α_T	Peak flux α_P	Fluence α_E
FD-SOC Theory	**3.0**	**2.33**	**2.0**	**1.67**	**1.50**
Lunar craters:					
Mare Tranquillitatis [1])	3.0				
Meteorites and debris [2])	2.75				
Asteroid belt:					
Spacewatch Surveys[3])	2.8				
Sloan Survey[4])	2.3-4.0				
Subaru Survey[5])	2.3				
Saturn ring:					
Voyager 1[6])	2.74-3.11				
Magnetosphere:					
Active substorms[7])		1.21 ± 0.08		1.05 ± 0.08	
Quiet substorms[7])		1.16 ± 0.03		1.00 ± 0.02	
Substorm flow bursts[8])			1.59 ± 0.07		
AE index bursts[9])			1.24		
AU index bursts[10])			1.3		
Outer radiation belt[11])				1.5-2.1	
Solar Flares:					
ISEE-3, HXR[12]			1.88-2.73	1.75-1.86	1.51-1.62
HXRBS/SMM, HXR[13]			2.17 ± 0.05	1.73 ± 0.01	1.53 ± 0.02
BATSE/CGRO, HXR[14]			2.20-2.42	1.67-1.69	1.56-1.58
RHESSI, HXR[15]			1.8-2.2	1.58-1.77	1.65-1.77
Yohkoh, SXR[16]	1.96-2.41	1.77-1.94		1.64-1.89	1.4-1.6
GOES, SXR[17]			2.0-5.0	1.86-1.98	1.88
SOHO/EIT, EUV[18]		2.3-2.6	1.4-2.0		
TRACE, EUV[19]	2.50-2.75	2.4-2.6		1.52-2.35	1.41-2.06
AIA/SDO, 335 A, EUV[20]	1.96		2.17	1.34	
Microwave bursts[21]				1.2-2.5	
Type III bursts[22]				1.26-1.91	
Solar energetic particles[23]				1.10-2.42	1.27-1.32
Stellar Flares:					
Flare stars (reported)[24]					2.17 ± 0.25
Flare stars (powerlaw fit)[24]					1.85 ± 0.31
Flare stars (cumulative fit)[24]					1.93 ± 0.35
Astrophysical Objects:					
Crab pulsar[25]				3.06-3.50	
PSR B1937+21[26]				2.8 ± 0.1	
Soft Gamma-Ray repeaters[27]					$1.43 - 1.76$
Cygnus X-1 black hole[28]				7.1	
Blazar GC 0109+224[29]				1.55	
Cosmic rays[30]					$2.7 - 3.3$

References to Table 13.2: [1]) Cross (1966); [2]) Sornette (2004); [3]) Jedicke and Metcalfe (1998); [4]) Ivezic et al. (2001); [5]) Yoshida et al. (2003), Yoshida and Nakamura (2007); [6]) Zebker et al. (1985), French and Nicholson (2000); [7]) Lui et al. (2000); [8]) Angelopoulos et al. (1999); [9]) Takalo (1993), Takalo et al. (1999); [10]) Freeman et al. (2000); Chapman and Watkins (2001); [11]) Crosby et al. (2005) [12]) Lu et al. (1993), Lee et al. (1993); [13]) Crosby et al. (1993); [14]) Aschwanden (2011a,b); [15]) Christe et al. (2008), Lin et al. (2001), Aschwanden (2011a,b); [16]) Shimizu (1995), Aschwanden and Parnell (2002); [17]) Lee et al. (1995), Feldman et al. (1997), Veronig et al. (2002a,b), Aschwanden and Freeland (2012); [18]) Krucker and Benz (1998), McIntosh and Gurman (2005); [19]) Parnell and Jupp (2000), Aschwanden et al. 2000, Benz and Krucker (2002), Aschwanden and Parnell (2002), Georgoulis et al. (2002); [20]) Aschwanden (2012b) [21]) Akabane (1956), Kundu (1965), Kakinuma et al. (1969), Das et al. (1997), Nita et al. (2002); [22]) Fitzenreiter et al. (1976), Aschwanden et al. (1995), Das et al. (1997), Nita et al. (2002); [23]) Van Hollebeke et al. (1975), Belovsky and Ochelkov (1979), Cliver et al. (1991), Gabriel and Feynman (1996), Smart and Shea (1997), Mendoza et al. (1997), Miroshnichenko et al. (2001), Gerontidou et al. (2002); [23]) Gabriel and Feynman (1996); [24]) Robinson et al. (1999), Audard et al. (2000), Kashyap et al. (2002), Güdel et al. (2003), Arzner and Güdel (2004), Arzner et al. (2007), Stelzer et al. (2007); [25]) Argyle and Gower (1972), Lundgren et al. (1995); [26]) Cognard et al. (1996); [27]) Gogus et al. (1999, 2000); [28]) Negoro et al. (1995), Mineshige and Negoro (1999); [29]) Ciprini et al. (2003); [30]) e.g., Fig. 13.18 (courtesy of Simon Swordy, Univ.Chicago).

Acknowledgements: This work has benefitted from fruitful discussions with Virginia Trimble, Shin Mineshige, Bob Stern, Manuel Güdel, comments from two anonymous referees, and by the *International Space Science Institute (ISSI)* at Bern Switzerland, which hosted and supported a workshop on *Self-Organized Criticality and Turbulence* during October 15-19, 2012.

13.4 References

Akabane, K. 1956, *Some features of solar radio bursts at around 3000 Mc/s*, Publ. Astron. Soc. Japan **8**, 173-181.

Angelopoulos, V., Mukai, T., and Kokubun, S. 1999, *Evidence for intermittency in Earth's plasma sheet and implications for self-organized criticality*, Phys. Plasmas **6**(11), 4161-4168.

Arzner, K. and Güdel, M. 2004, *Are coronae of magnetically active stars heated by flares? III. Analytical distribution of superposed flares*, Astrophys. J. **602**, 363-376.

Arzner, K., Güdel, M., Briggs, K., Telleschi, A., and Audard, M. 2007, *Statistics of superimposed flares in the Taurus molecular cloud*, Astron. Astrophys. **468**, 477-484.

Aschwanden, M.J., Benz, A.O. and Schwartz, R.A. 1993, *The timing of electron beam signatures in hard X-rays and radio: solar flare observations*

with BATSE/Compton Gamma-Ray Observatory and PHOENIX, Astrophys. J. **417**, 790-804.

Aschwanden, M.J., Benz, A.O., Dennis, B.R., and Schwartz, R.A. 1995, *Solar electron beams detected in hard X-rays and radio waves*, Astrophys. J. **455**, 347-365.

Aschwanden, M.J., Tarbell, T., Nightingale, R., Schrijver, C.J., Title, A., Kankelborg, C.C., Martens, P.C.H., and Warren, H.P. 2000, *Time variability of the quiet Sun observed with TRACE: II. Physical parameters, temperature evolution, and energetics of EUV nanoflares*, Astrophys. J. **535**, 1047-1065.

Aschwanden, M.J. 1999, *Do EUV nanoflares account for coronal heating?*, Solar Phys. **190**, 233-247.

Aschwanden, M.J. and Parnell, C.E. 2002, *Nanoflare statistics from first principles: fractal geometry and temperature synthesis*, Astrophys. J. **572**, 1048-1071.

Aschwanden, M.J. 2011a, *Self-Organized Criticality in Astrophysics. The Statistics of Nonlinear Processes in the Universe*, ISBN 978-3-642-15000-5, Springer-Praxis: New York, 416p.

Aschwanden, M.J. 2011b, *The state of self-organized criticality of the Sun during the last three solar cycles. I. Observations*, Solar Phys **274**, 99-117.

Aschwanden,M.J. 2012a, *A statistical fractal-diffusive avalanche model of a slowly-driven self-organized criticality system*, Astron.Astrophys. **539**, A2, (15 p).

Aschwanden,M.J. 2012b, *The spatio-temporal evolution of solar flares observed with AIA/SDO: Fractal diffusion, sub-diffusion, or logistic growth ?* Astrophys. J. **757**, 94.

Aschwanden,M.J. and Freeland,S.L. 2012, *Automated solar flare statistics in soft X-rays over 37 years of GOES observations - The invariance of self-organized criticality during three solar cycles*, Astrophys. J. **754**, 112.

Audard, M., Güdel, M., Drake, J.J., and Kashyap, V.L. 2000, *Extreme-ultraviolet flare activity in late-type stars* Astrophys. J. **541**, 396-409.

Bai, T. 1993, *Variability of the occurrence frequency of solar flares as a function of peak hard X-ray rate*, Astrophys. J. **404**, 805-809.

Belovsky, M.N., and Ochelkov, Yu. P. 1979, *Some features of solar-flare electromagnetic and corpuscular radiation production*, Izvestiya AN SSR, Phys. Ser. **43**, 749-752.

Benz, A.O. and Krucker, S. 2002, *Energy distribution of microevents in the quiet solar corona*, Astrophys. J. **568**, 413-421.

Biesecker, D.A. 1994, *On the occurrence of solar flares observed with the Burst and Transient Source Experiment (BATSE)*, PhD Thesis, University of New Hampshire.

Chapman, S.C. and Watkins, N. 2001, *Avalanching and self-organised criticality, a paradigm for geomagnetic activity?*, Space Sci. Rev. **95**, 293-307.

Christe, S., Hannah, I.G., Krucker, S., McTiernan, J., and Lin, R.P. 2008, *RHESSI microflare statistics. I. Flare-finding and frequency distributions*, Astrophys. J. **677**, 1385-1394.

Ciprini, S., Fiorucci, M., Tosti, G., and Marchili, N. 2003, *The optical variability of the blazar GV 0109+224. Hints of self-organized criticality*, in *High energy blazar astronomy*, ASP Conf. Proc. **229**, (eds. L.O. Takalo and E. Valtaoja), ASP: San Francisco, p.265.

Cliver, E., Reames, D., Kahler, S., and Cane, H. 1991, *Size distribution of solar energetic particle events*, Internat. Cosmic Ray Conf. 22nd, Dublin, LEAC A92-36806 15-93, NASA:Greenbelt, p. 2:1-4.

Crosby, N.B., Aschwanden, M.J., and Dennis, B.R. 1993, *Frequency distributions and correlations of solar X-ray flare parameters*, Solar Phys. **143**, 275-299.

Crosby, N.B., Meredith, N.P., Coates, A.J., and Iles, R.H.A. 2005, *Modelling the outer radiation belt as a complex system in a self-organised critical state*, Nonlinear Processes in Geophysics **12**, 993-1001.

Cross, C.A. 1966, *The size distribution of lunar craters*, MNRAS **134**, 245-252.

Das, T.K., Tarafdar, G., and Sen, A.K. 1997, *Validity of power law for the distribution of intensity of radio bursts*, Solar Phys. **176**, 181-184.

Feldman, U., Doschek, G.A., and Klimchuk, J.A. 1997, *The occurrence rate of soft X-ray flares as a function of solar activity*, Astrophys. J. **474**, 511-517.

Fitzenreiter, R.J., Fainberg, J., and Bundy, R.B. 1976, *Directivity of low frequency solar type III radio bursts*, Solar Phys. **46**, 465-473.

Freeman, M.P., Watkins, N.W., and Riley, D.J. 2000, *Evidence for a solar wind origin of the power law burst lifetime distribution of the AE indices*, Geophys. Res. Lett. **27**, 1087-1090.

French, R.G. and Nicholson, P.D. 2000, *Saturn's rings. II. Partice sizes inferred from stellar occultation data*, Icarus **145**, 502-523.

Gabriel, S.B. and Feynman, J. 1996, *Power-law distribution for solar energetic proton events*, Solar Phys. **165**, 337-346.

Georgoulis, M.K., Kluivin,R. and Vlahos,L. 1995, *Extended instability criteria in isotropic and anisotropic energy avalanches*, Physica A **218**, 191-213.

Gerontidou, M., Vassilaki, A., Mavromichalaki, H., and Kurt, V. 2002, *Frequency distributions of solar proton events*, J. Atmos. Solar-Terr. Physics **64**(5-6), 489-496.

Gogus, E., Woods, P.M., Kouveliotou, C., van Paradijs, J., Briggs, M.S., Duncan, R.C., and Thompson, C. 1999, *Statistical properties of SGR 1900+14 bursts*, Astrophys. J. **526**, L93-L96.

Gogus, E., Woods, P.M., Kouveliotou, C., and van Paradijs, J. 2000, *Statistical properties of SGR 1806-20 bursts*, Astrophys. J. **532**, L121-L124.

Greenberg,R., Davies, D.R., Harmann, W.K., and Chapman, C.R., 1977, Icarus 30, 769-779.

Güdel, M., Audard, M., Kashyap, V.L., and Guinan, E.F. 2003, *Are coronae of magnetically active stars heated by flares? II. Extreme Ultraviolet and X-ray flare statistics and the differential emission measure distribution*, Astrophys. J. **582**, 423-442.

Güdel, M. 2004, *X-ray astronomy of stellar coronae*, Astron. Astrophys. Rev. **12**(2-3), 71-237.

Hudson, H.S. 1991, *Solar flares, microflares, nanoflares, and coronal heating*, Solar Phys. **133**, 357-369.

Ivezic, Z., Tabachnik, S., Rafikov, R., Lupton, R.H., Quinn, T., Hammergren, M., Eyer, L., Chu, J., Armstrong, J.C., Fan, X., Finlator, K., Geballe, T.R., Gunn, J.E., Hennessy, G.S., Knapp, G.R., et al. (SDSS Collaboration) 2001, *Solar system objects observed in the Sloan Digital Sky Survey Commissioning Data*, Astronomical J. **122**, 2749-2784.

Jedicke, R. and Metcalfe, T.S. 1998, *The orbital and absolute magnitude distributions of main belt asteroids*, Icarus **131**(2), 245-260.

Kakinuma, T., Yamashita, T., and Enome, S. 1969, *A statistical study of solar radio bursts a microwave frequencies*, Proc. Res. Inst. Atmos. Nagoya Univ. Japan, Vol. **16**, 127-141.

Kashyap, V.L., Drake, J.J., Güdel, M., and Audard, M. 2002, *Flare heating in stellar coronae*, Astrophys. J. **580**, 1118-1132.

Kouveliotou,C., Dieters, S., Strohmayer, T., van Paradijs, J., Fishman, G.J., Meegan, C.A., Hurley, K., Kommers, J., Smith, I., Frail, D., Muakami, T. 1998, *An X-ray pulsar with a superstrong magnetic field in the soft γ-ray repeater SGR 1806-20*, Nature **393**, 235-237.

Kouveliotou,C., Strohmayer, T., Hurley, K., van Paradijs, J., Finger, M.H., Dieters, S., Woods, P., Thomson, C., and Duncan, R.C. 1999, *Discovery of a magnetar associated with the soft gamma ray repeater SGR 1900+14*, Astrophys. J. **510**, L115-L118.

Krucker, S. and Benz, A.O. 1998, *Energy distribution of heating processes in the quiet solar corona*, Astrophys. J. **501**, L213-L216.

Kundu, M.R. 1965, *Solar radio astronomy*, Interscience Publication: New York, 660 p.

Lee, T.T., Petrosian, V., and McTiernan, J.M. 1993, *The distribution of flare parameters and implications for coronal heating*, Astrophys. J. **412**, 401-409.

Lee, T.T., Petrosian, V., and McTiernan, J.M. 1995, *The Neupert effect and the chromospheric evaporation model for solar flares*, Astrophys. J. **418**, 915-924.

Lin, R.P., Feffer,P.T., and Schwartz,R.A. 2001, *Solar Hard X-Ray Bursts and Electron Acceleration Down to 8 keV*, Astrophys. J. **557**, L125-L128.

Lu, E.T. and Hamilton, R.J. 1991, *Avalanches and the distribution of solar flares*, Astrophys. J. **380**, L89-L92.

Lu, E.T., Hamilton, R.J., McTiernan, J.M., and Bromund, K.R. 1993, *Solar flares and avalanches in driven dissipative systems*, Astrophys. J. **412**, 841-852.

Lui, A.T.Y., Chapman, S.C., Liou,K., Newell, P.T., Meng, C.I., Brittnacher, M., and Parks, G.K. 2000, *Is the dynamic magnetosphere an avalanching system?*, Geophys. Res. Lett. **27**(7), 911-914.

McIntosh, S.W. and Gurman, J.B. 2005, *Nine years of EUV bright points*, Solar Phys. **228**, 285-299.

Mendoza, B., Melendez-Venancio, R., Miroshnichenko, L.I., and Perez-Enriquez, R. 1997, *Frequency distributions of solar proton events*, Proc. 25th Int. Cosmic Ray Conf. **1**, 81.

Melatos, A., Peralta, C., and Wyithe, J.S.B. 2008, *Avalanche Dynamics of radio pulsar glitches*, Astrophys. J. **672**, 1103-1118.

Melatos, A. and Warszawski, L., 2008, *Superfluid vortex ynpinning as a coherent noise process, and the scale invariance of pulsar glitches*, Astrophys. J. **700**(2), 1254-1540.

Mineshige, S. and Negoro, H. 1999, *Accretion disks in the context of self-organized criticality: How to produce 1/f fluctuations ?*, in *High energy processes in accreting black holes*, ASP Conf. Ser. **161**, 113-128.

Miroshnichenko, L.I., Mendoza, B., and Perez-Enriquez R. 2001, *Size distributions of the >10 MeV solar proton events*, Solar Phys. **202**, 151-171.

Nishizuka, N., Asai, A., Takasaki, H., Kurokawa, H., and Shibata, K. 2009, *The Power-Law Distribution of Flare Kernels and Fractal Current Sheets in a Solar Flare*, Astrophys. J. **694**, L74-L77.

Negoro, H., Kitamoto, S., Takeuchi, M., and Mineshige, S. 1995, *Statistics of X-ray fluctuations from Cygnus X-1: Reservoirs in the disk ?* Astrophys. J. **452**, L49-L52.

Nita, G.M., Gary, D.E., Lanzerotti, L.J., and Thomson, D.J. 2002, *The peak flux distribution of solar radio bursts*, Astrophys. J. **570**, 423-438.

Parnell,C.E. and Jupp,P.E. 2000, *Statistical analysis of the energy distribution of nanoflares in the quiet Sun* Astrophys. J. **529**, 554-569.

Robinson, R.D., Carpenter, K.G., and Percival, J.W. 1999, *A search for microflaring activity on dMe flare stars. II. Observations of YZ Canis Minoris*, Astrophys. J. **516**, 916-923.

Rosner, R., and Vaiana, G.S. 1978, *Cosmic flare transients: constraints upon models for energy storage and release derived from the event frequency distribution*, Astrophys. J. **222**, 1104-1108.

Shibata, K. and Yokoyama T. 1999, *Origin of the universal correlation between the flare temperature and the emission measure for solar and stellar flares*, Astrophys. J. **526**, L49-L52.

Shibata, K. and Yokoyama T. 2002, *A Hertzsprung-Russell-like diagram for solar/stellar flares and corona: emission measure versus temperature diagram*, Astrophys. J. **577**, 422-432.

Shimizu, T. 1995, *Energetics and occurrence rate of active-region transient brightenings and implications for the heating of the active-region corona*, Publ. Astron. Soc. Japan **47**, 251-263.

Smart, D.F. and Shea, M.A. 1997, *Comment on the use of solar proton spectra in solar proton dose calculations*, in Proc. *Solar-Terrestrial Prediction Workshop V*, Hiraiso Solar-Terrestrial Research Center, Japan, p.449.

Sornette, D. 2004, *Critical phenomena in natural sciences: chaos, fractals, self-organization and disorder: concepts and tools*, Springer, Heidelberg, 528 p.

Stelzer, B., Flaccomio, E., Briggs, K., Micela, G., Scelsi, L, Audard, M., Pillitteri, I., and Güdel, M. 2007, *A statistical analysis of X-ray variability in pre-main sequence objects of the Taurus molecular cloud*, Astron. Astrophys. **468**, 463-475.

Takalo, J. 1993, *Correlation dimension of AE data*, Ph. Lic. Thesis, *Laboratory report 3*, Dept. Physics, University of Jyväskylä.

Takalo, J., Timonem, J., Klimas, A., Valdivia, J., and Vassiliadis, D. 1999, *Nonlinear energy dissipation in a cellular automaton magnetotail field model* Geophys. Res. Lett. **26**(13), 1813-1816.

Takeuchi, M., Mineshige, S., and Negoro, H. 1995, *X-ray fluctuations from black-hole objects and self organization of accretion disks*, Publ. Astron. Soc. Japan **47**, 617-627.

Thompson, C. and Duncan, R.C. 1996, *The soft gamma repeaters as very strongly magnetized neutron stars. II. Quiescent neutrino, X-ray, and Alfvén wave emission* Astrophys. J. **473**, 322-342.

VanHollebeke, M.A.I., Ma Sung L.S., and McDonald F.B. 1975, *The variation of solar proton energy spectra and size distribution with heliolongitude*, Solar Phys. **41**, 189-223.

Veronig, A., Temmer, M.,Hanslmeier, A., Otruba, W., and Messerotti, M. 2002a, *Temporal aspects and frequency distributions of solar X-ray flares*, Astron. Astrophys. **382**, 1070-1080.

Veronig, A., Temmer, M., and Hanslmeier, A. 2002b, *Frequency distributions of solar flares*, Hvar Observatory Bulletin **26**(1), 7-12.

Warszawski, L, and Melatos, A. 2008, *A cellular automaton model of pulsar glitches*, MNRAS **390**(1), 175-191.

Warszawski, L, and Melatos, A. 2012, *Knock-on processes in superfluid vortex avalanches and pulsar glitdch statistics*, MNRAS (in press), eprint arXiv:1210.2203.

Yoshida, F., Nakamura, T., Watanab, J., Kinoshita, D., and Yamamoto, N., 2003, *Size and spatial distributions of sub-km main-belt asteroids*, Publ. Astron. Soc. Japan **55**, 701-715.

Yoshida, F. and Nakamura T. 2007, *Subary Main Belt Asteroid Survey (SMBAS) - Size and color distributions of small main-belt asteroids*, Planet. Space Science **55**, 113-1125.

Zebker, H.A., Marouf, E.A., and Tyler, G.L. 1985, *Saturn's rings particle size distributions for a thin layer model*, Ikarus **64**, 531-548.