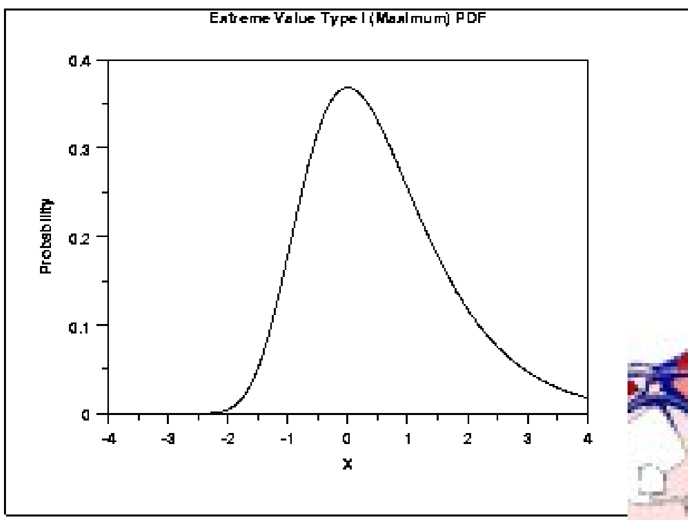
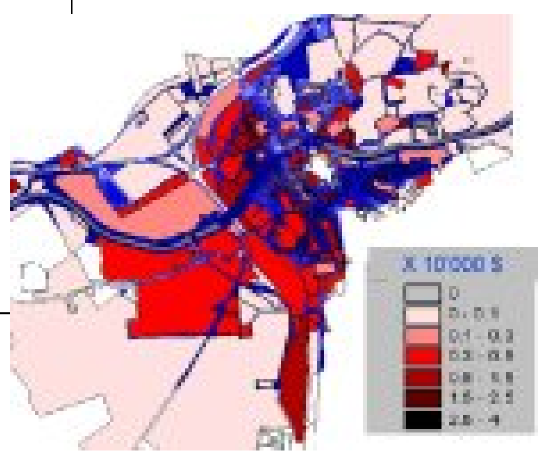
Extreme value statistics
and
slow dynamics of random systems

Valerii Vinokur

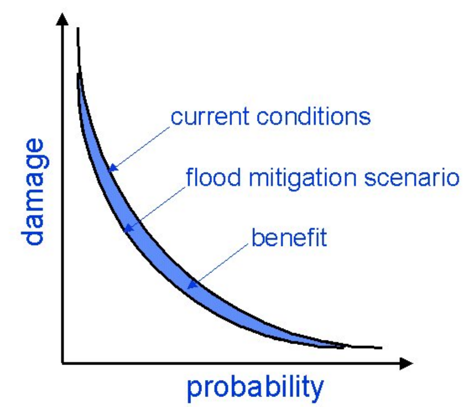
Argonne National Laboratory



Flood damage distribution



Collaborations:
M.C. Marchetti
S. Scheidl

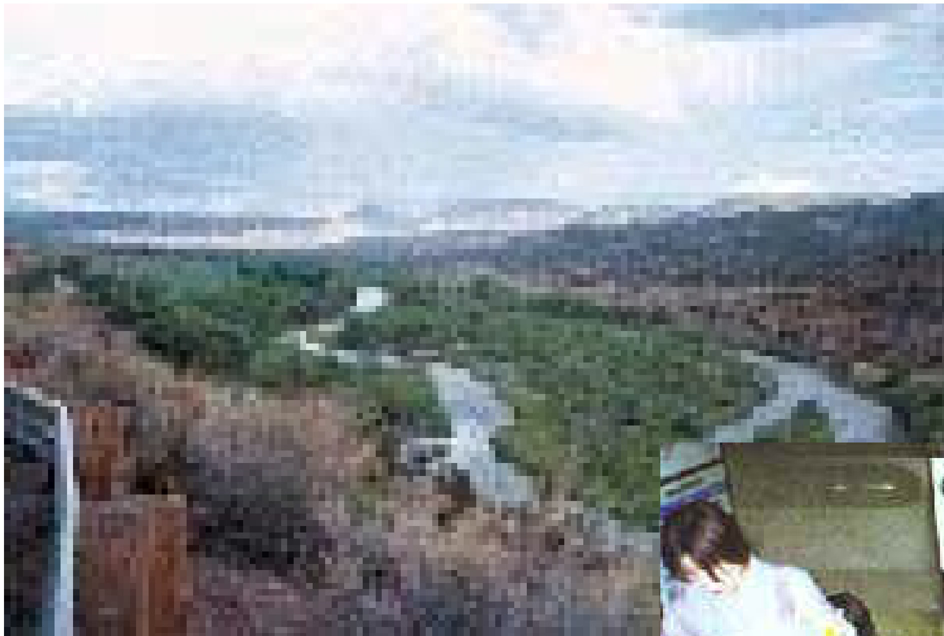


Outline

1. *Problems of real life…and history*
2. *Extreme value statistics*
3. *Dynamics of exemplary random system: domain wall, etc…*
4. *Barriers distribution*

The oldest problems connected with extreme values arise from floods

Their economic importance was early realized since ancient agrarian economies were based exclusively on water flow



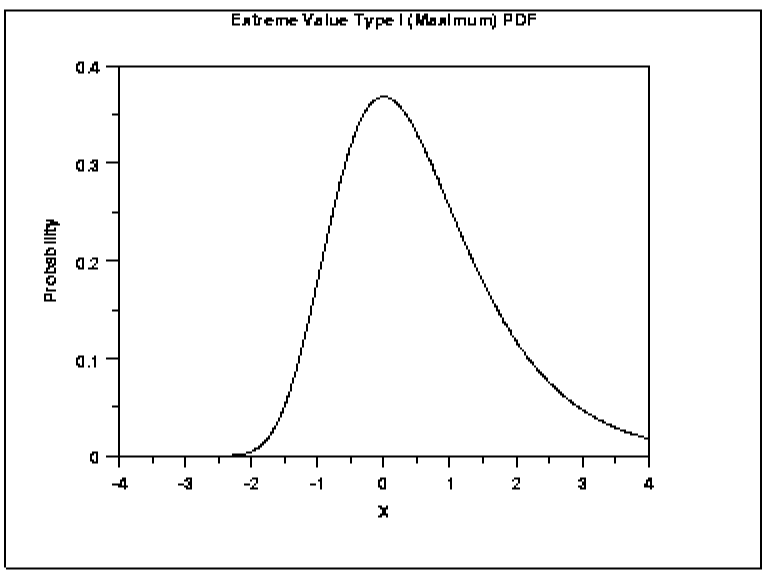
Their importance has increased
In industrial economies

An inundation happens when water flows where it ought not to flow

Consider the mean daily discharge of a river at a specific station.

Among the 365 daily discharges during a year, there is one measure which is the largest. This discharge is called annual flood.

The assessment of risks and damage associated with floods is crucial.



The critical question is how to find the right distribution of floods

History…

1909. Nicolaus Bernoulli: *n* men of equal age die within *t* years. What is the mean duration of life of the last survivor?

1925. L.H.C. Tippett: probabilities of the largest normal values for different sample sizes up to 1000

1927. M. Fréchet introduced stability postulate

The basic work for extreme values:

R.A. Fisher and L.H.C. Tippett, Proc. Cambridge Phil. Soc. **24**, 180 (1928).

B.V. Gnedenko and A.N. Kolmogorov, *Limit distributions for Sums of Independent Random Variables* (Addison Wesley, Reading, MA, 1954).

E.J. Gumbel, *Statistical Theory of Extreme Values and Some Practical Applications*, National Bureau of Standards Applied Mathematics Series, vol. 33 (1954); *Statistics of Extremes* (Columbia University Press, New York, 1958).

1. Extreme value statistics

1.Extreme statistics: What is the asymptotic distribution of the maximal (or minimal) value reached by the variable X?
How does this distribution depend on the distribution of random force?

2.Persistence probability: What is the probability that the variable has never attained a certain threshold?

3.Large deviations: What is the probability that the variable has assumed a positive sign for a certain fraction of the total time elapsed?
Is this distribution peaked around its mean value or not?

4.Persistent large deviations (Generalized persistence): What is the probability that the fraction of time during which the variable has been positive, has always been larger than a certain value?

[Fatigue analysis]

Given: a set of N random variables $\{X_i\}$

Example: imagine a complex device, containing a number of critical components each one characterized by a life time X_i

The life time of the whole device is determined by the minimum between component lifetimes. In such cases the useful characteristics of the random sample are their extremes:

$$L_n = \min_{i=1,2,\dots n} \{X_i\} \qquad W_n = \max_{i=1,2,\dots n} \{X_i\}$$

We are interested in the limit of large samples $n \gg 1$, and, in particular, in the asymptotic distribution of the random descriptor W_n (sum, average, minimum, maximum or whatever else).

Rigorously the problem can be formulated in the following manner:

We have to determine two sequences a_n and b_n
and a limit distribution $H(x)$ such that, when $n \to \infty$:

$$\Pr\mathrm{ob}\left[\frac{W_n - a_n}{b_n} \le x\right] \to H(x)$$

The main known results in probability theory are for independent identically distributed random variables (hereafter abbreviated as **iid** variables), i.e. when the joint probability distribution of the sample is simply:

$$\Pr\mathrm{ob}[X_1 \le x_1, \ldots, X_n \le x_n] = \Pr\mathrm{ob}[X_1 \le x_1] \ldots \Pr\mathrm{ob}[X_n \le x_n]$$

and all the X's have the same distribution $\Pr\mathrm{ob}[X_i \le x] = F(x)$

The problem of extreme statistics of iid variables can be easily cast:
if $F(x)$ is the distribution of each iid variable forming the sample,
then the probability distribution for the maximal value $W_n = \max_{i=1,2,\ldots n} \{X_i\}$ is:

$$\mathrm{Prob}[W_n \leq x] = \mathrm{Prob}[X_1 \leq x, \ldots, X_n \leq x]$$
$$= [F(x)]^n \equiv H_n(x)$$

where the last equality is the definition of $H_n(x)$

The asymptotic distribution problem, hence, is the determination
of a_n, b_n, and $H_n(x)$ such that:

$$\lim_{n\to\infty} (F(a_n + b_n x))^n = \lim_{n\to\infty} H_n(a_n + b_n x) = H(x)$$

While the exact solution for the problem of the asymptotic distribution of the sum of iid random variables is widely known, the relation between extreme statistics and statistical physics has not been deeply investigated. That such a relation should exist is however intuitively obvious: for example, at low temperature a disordered system will preferentially occupy its low-energy (i.e. minimal) states, which are random variables because of the disordered nature of the problem.

Extreme Statistics

$$\{X_1,\ldots,X_n\} \to F(X_1,\ldots,X_n)$$

- Extreme value statistics after n trials (ex. minimum W_n):
  - "typical" value (a_n)
  - "typical" fluctuations (b_n)
  - reduced variable
  - asymptotic distribution

$$z_n = \frac{W_n - a_n}{b_n}$$

$$P(z_n) \xrightarrow[n\to\infty]{} H(z)$$

Independent an identically distributed random variables

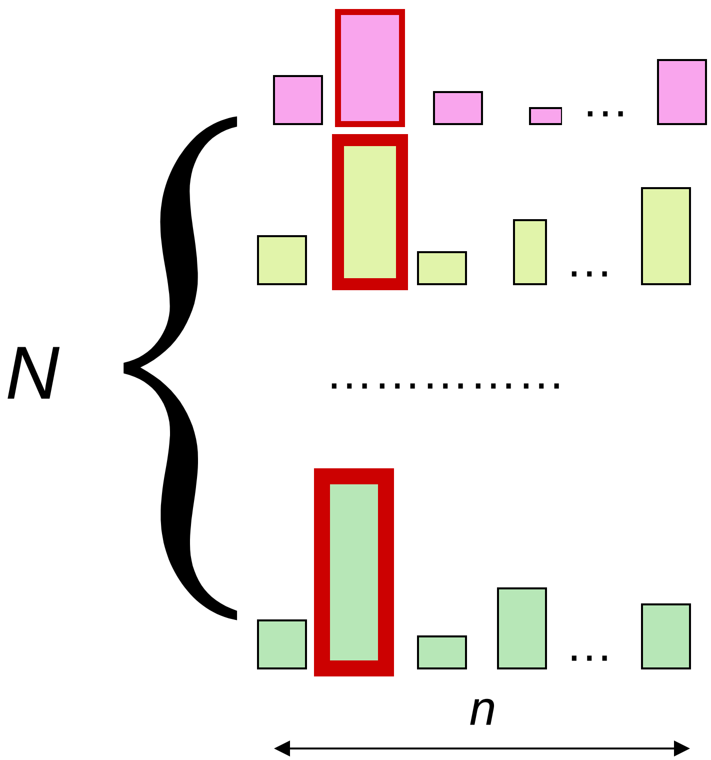
$$F(X_1,\ldots,X_n) = f(X_1)\ldots f(X_n)$$

Three universality classes for $H(z)$, depending from the tail of $f(x)$

- **Gumbel (exponential tail of $f(x)$)**
- Weibull (bounded variables)
- Fréchet (power law tail of $f(x)$)

The Stability Postulate

Consider *N* samples, each of size *n* –



The distribution of the largest value in *Nn* observations will tend to the same asymptotic expression as the distribution of the largest value in samples of size *n*, provided that such an asymptote exists.

For the probability that the largest value is below *x*:

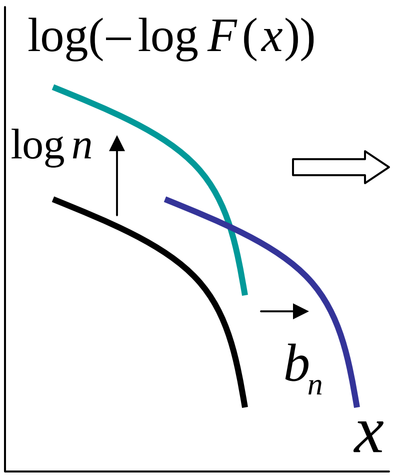
$$F^n(x) = F(a_n x + b_n)$$

Stability postulate

$$F^{n}(x) = F(a_n x + b_n)$$

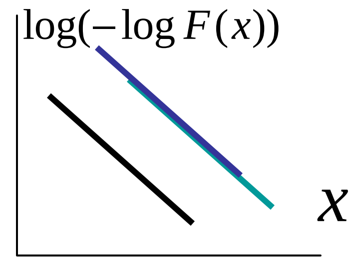
1. $a_n = 1$

$$F^{n}(x) = F(x + b_n)$$



$$\log n + \log(-\log F(x)) = \log(-\log F(x + a_n))$$

$$\Longrightarrow \log(-\log F(x)) = Const + \frac{x \log n}{b_n}$$



$$F^{nm}(x) = F(x + b_{nm})$$

$$[F^{n}(x)]^{m} = F(x + b_n + b_m)$$

$$b_n + b_m = b_{nm} \Longrightarrow \boxed{b_n = c \log n}$$

$$\log[-\log F(x)] = x / c + k \Longrightarrow$$

$$-\log F(x) = e^{-\alpha(x-u)}, \alpha = -1/c > 0 \quad \boxed{F(x) = \exp(-ne^{-\alpha x})}$$

2. If an differs from unity the two curves $F^n(x)$ and $F(x)$ are no longer parallel. There is a value x' *where the two* probabilities are equal. Now the equation $F^n(x') = F(x')$ can be satisfied if and only if $F(x')=0$ or $F(x') = 1$

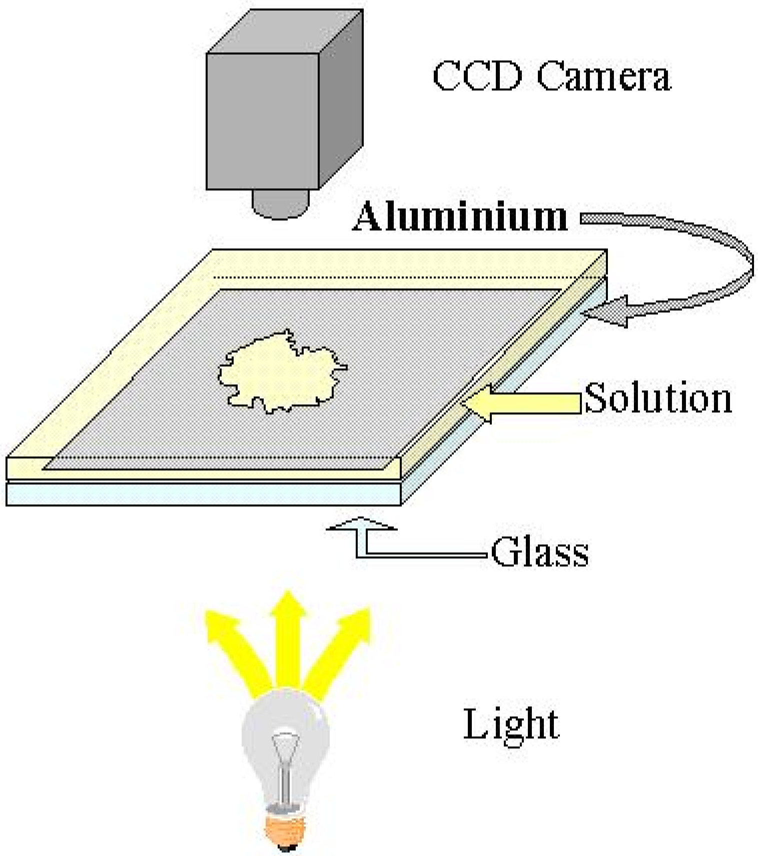
In the first (second) case the distribution starts (ends) with x' and is bounded to the left (right)

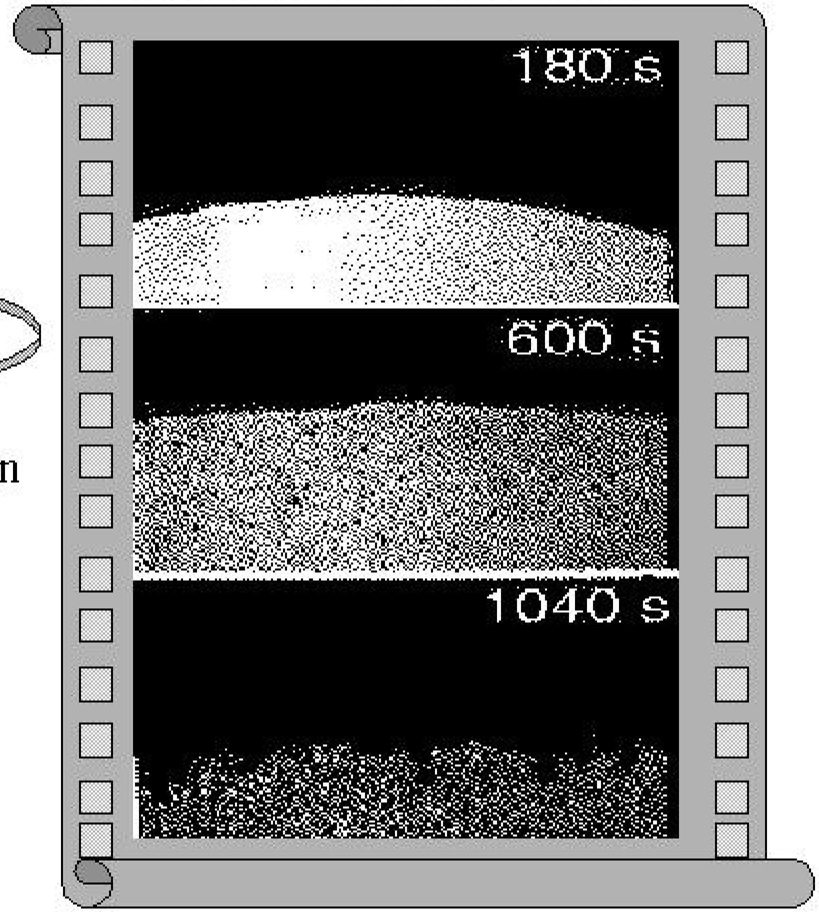
$$\boxed{F(x) = \exp(-ne^{-\alpha x})}$$

Example of application in physics:

Etching of a thin film of aluminium immersed in a corrosive solution

Experiment: *L.Balazs et F.Gouyet (1995), L.Balazs (1996)* PMC Ecole Polytechnique.





Etching model

B. Sapoval et al. (1998)

$N_{et}(t)$: Number of etching "molecules"

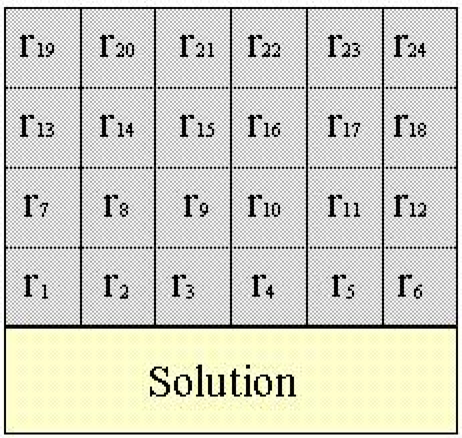
V: Volume of the corrosive solution

$$p(t) = \frac{N_{et}(t)}{V}$$

Corrosion power of the solution at time *t*

t=0

Solid



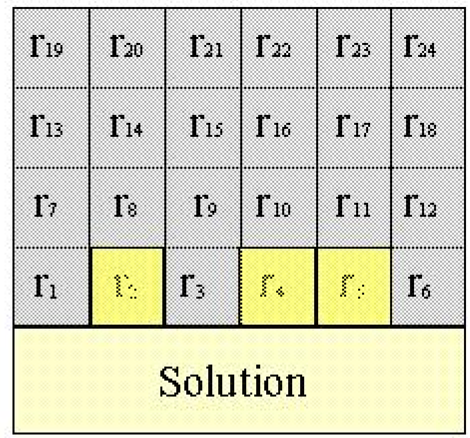
$$r_2, r_4, r_5 < p(0)$$

$$N_{et}(1) = N_{et}(0) - 3$$

three solid « sites » corroded

t=1

Solid



$$p(0) = \frac{N_{et}(0)}{V}$$

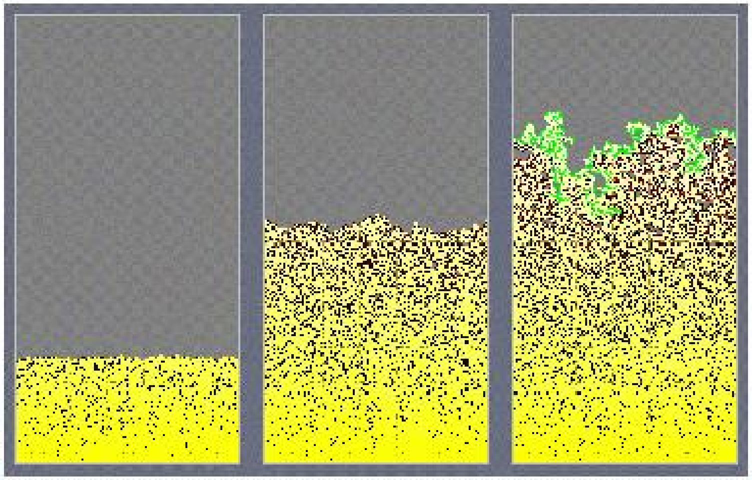
$$p(1) = \frac{N_{et}(1)}{V} = \frac{N_{et}(0) - 3}{V}$$

Study of the model: phenomenology

In collaboration with
B.Sapoval & A.Gabrielli

Numerical simulations

- ***L*: Size of the attacked side;**
- ***V*: volume of the etching solution;**
- *p(0)*: starting value of the corrosive power fixed to 1 ($N_{et} = p(0)\, V$)

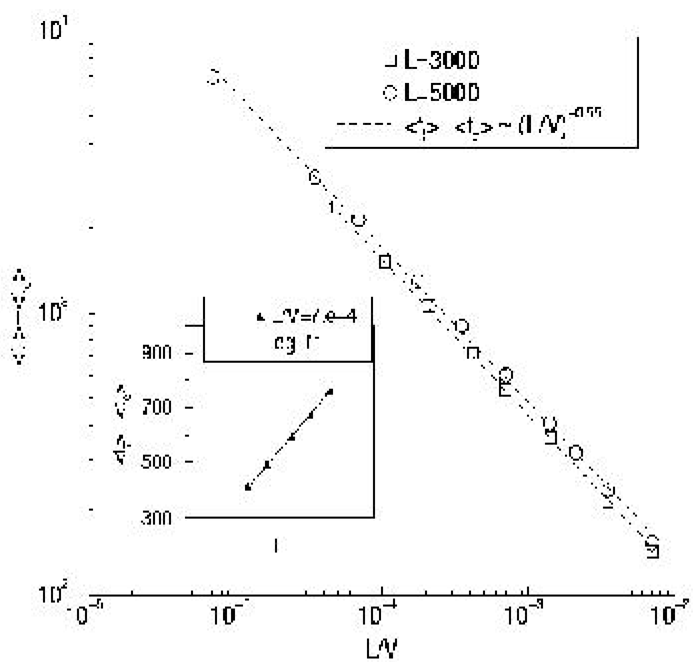


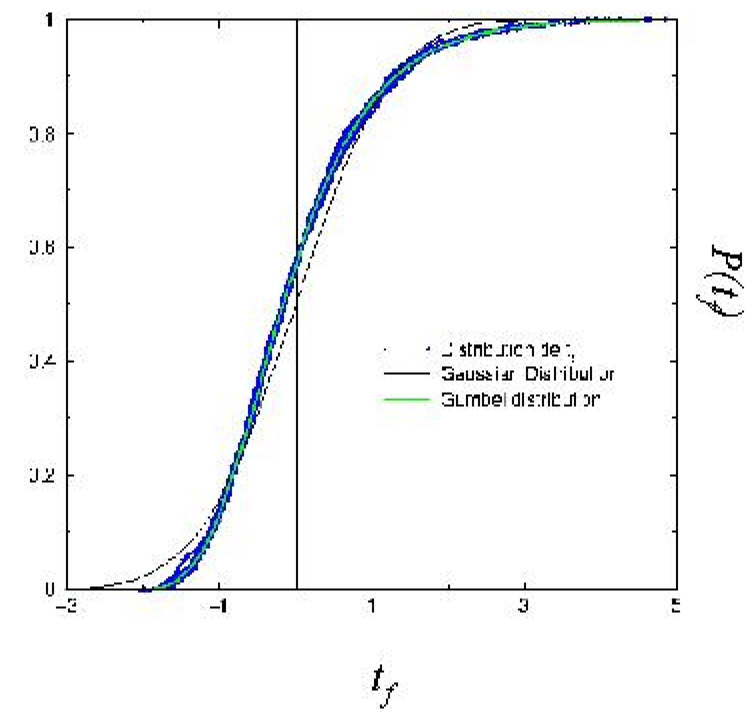
- **Etching dynamics:**
  - *First regime: smooth interface liquid-solid*
  - *Production of debris of growing size*
  - *Growing roughness of the interface*
- **Spontaneous arrest of dynamics at time t_f:**
  - *Final fractal interface liquid-solid (geometric correlations)*

Quantity extremal: *arrest time of the process*

(same behavior for the *maximal depth* attained by the solution inside the solid)

Corrections to the scaling behavior ⟷ Gumbel distribution of fluctuations

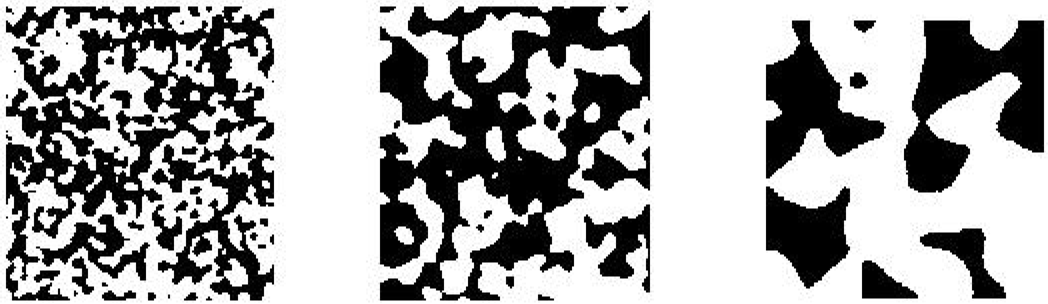




Domain growth

It seems self-similar...

- Quench from the disordered to the ordered phase (Temperature: *T* = *infinity* -> *0*



t=200 t=800 t=3200

- Domain growth: diverging characteristic length

$$L(t) \approx t^{1/z}$$

Universal dynamical exponent

it seems simple!

- Conserved order parameter: z=3, **not conserved: z=2**

Persistence

What is the probability for a spin to never flip up to time *t* ?

It is not so simple

$$R(t) \approx t^{-\theta}$$

Persistence exponent

- Ising model — 1d: θ=3/8 (*analytique*); 2d: θ=0.22 (*numérique*)
- Diffusion Equation — 1d: θ=0.121 (*numérique*); 2d: θ=0.188 (*numérique*)

$$\left\{ \begin{aligned} s_i(t+1) &= \mathrm{sign}\Big(\sum_{<j>} s_j(t)\Big) \\ \sigma(t) &= s_{i_0}(t) \end{aligned} \right\}$$

$$\left\{ \begin{aligned} \frac{\partial \varphi(x,t)}{\partial t} &= \nabla^2 \varphi(x,t) \\ \sigma(t) &= \mathrm{sign}\,(\varphi(x_0,t)) \end{aligned} \right\}$$

It depends on model, dimensionality, dynamics details, anisotropy, ...etc..

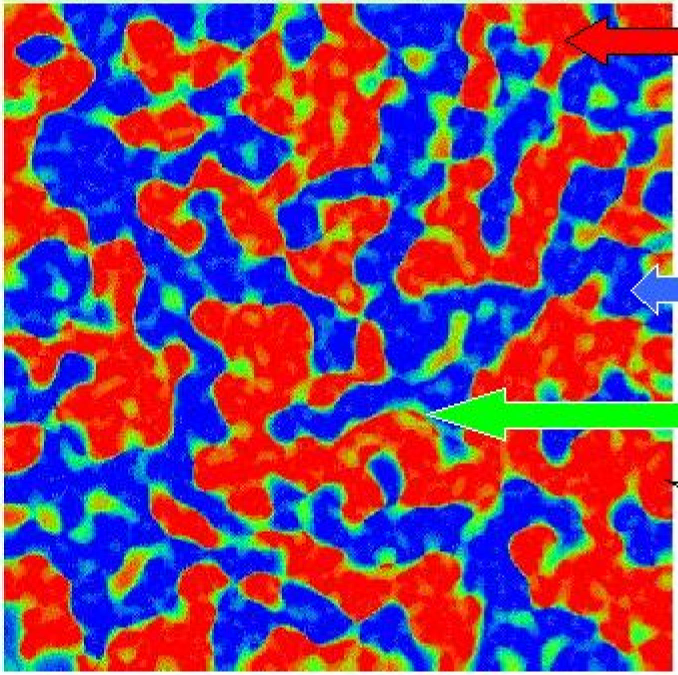
Occupation times and large deviations

I.Dornic et C.Godrèche (1998)

- Time averaged magnetisation

$$M(t) \equiv \frac{1}{t}\int_0^t \sigma(u)\,du$$

Example: 2d Ising Model



Red Spins:

- always ~ *+1, i.e.* (« almost ») persistent *M(t) ~ 1*

Blue Spins:

- always ~ *-1, i.e. M(t) ~ -1*

Green Spins:

- M(t) ~ *0*

Average value is not the most probable value

Persistence Spectrum

H.Chaté, A.Lemaître

$$R(t,x) \rightarrow t^{-\theta(x)}$$

$\theta(1)=\theta$ (Usual) Persistence exponent

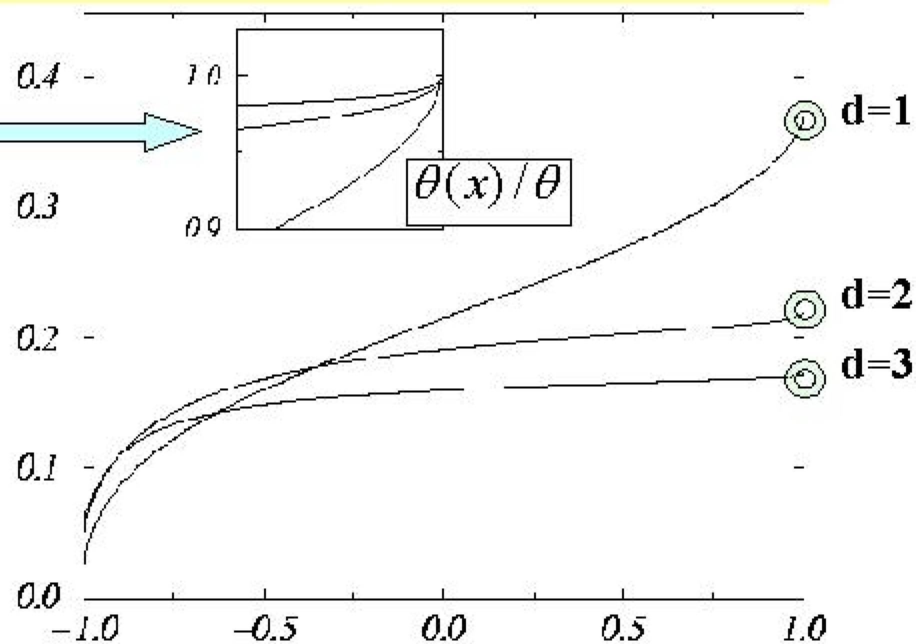
$\theta(-1)=0$ (because $R(t,-1)=1$)

For $-1<x<1$, family of exponents?

Spectrum (numerical) Ising model d=1,2,3

« Singular» behavior near $x \sim 1$

(but decreasing for larger dimensions)



Usual persistent exponents

Exemplary system: elastic medium in random environment.
Models a wealth of physical systems and phenomena:

$$\mathcal{H} = \int d^D x \left[\frac{C}{2} \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)^2 + V(\mathbf{x}, \mathbf{u}) - \mathbf{F} \cdot \mathbf{u} \right]$$

Dislocations in crystals — *CDW and SDW*

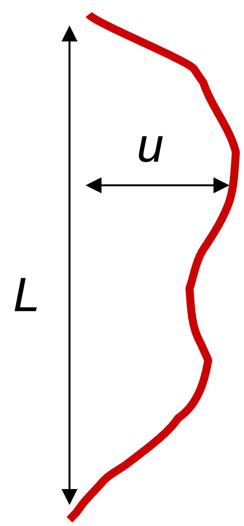
Domain walls — *Interacting electrons*

Wigner crystals on disordered substrates — *Spin- and other glasses…*

$$\mathcal{H} = \int d^D x \left[\frac{C}{2} \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} \right)^2 + V(\mathbf{x}, \mathbf{u}) - \mathbf{F} \cdot \mathbf{u} \right]$$

$$\langle V(\mathbf{x}, \mathbf{u}) V(\mathbf{x}', \mathbf{u}') \rangle = \Delta^2 \delta^D (\mathbf{x} - \mathbf{x}') f(|\mathbf{u} - \mathbf{u}'|/\xi)$$

$$w(L) = \langle [\mathbf{u}(\mathbf{x} + \mathbf{L}) - \mathbf{u}(\mathbf{x})]^2 \rangle^{1/2}$$

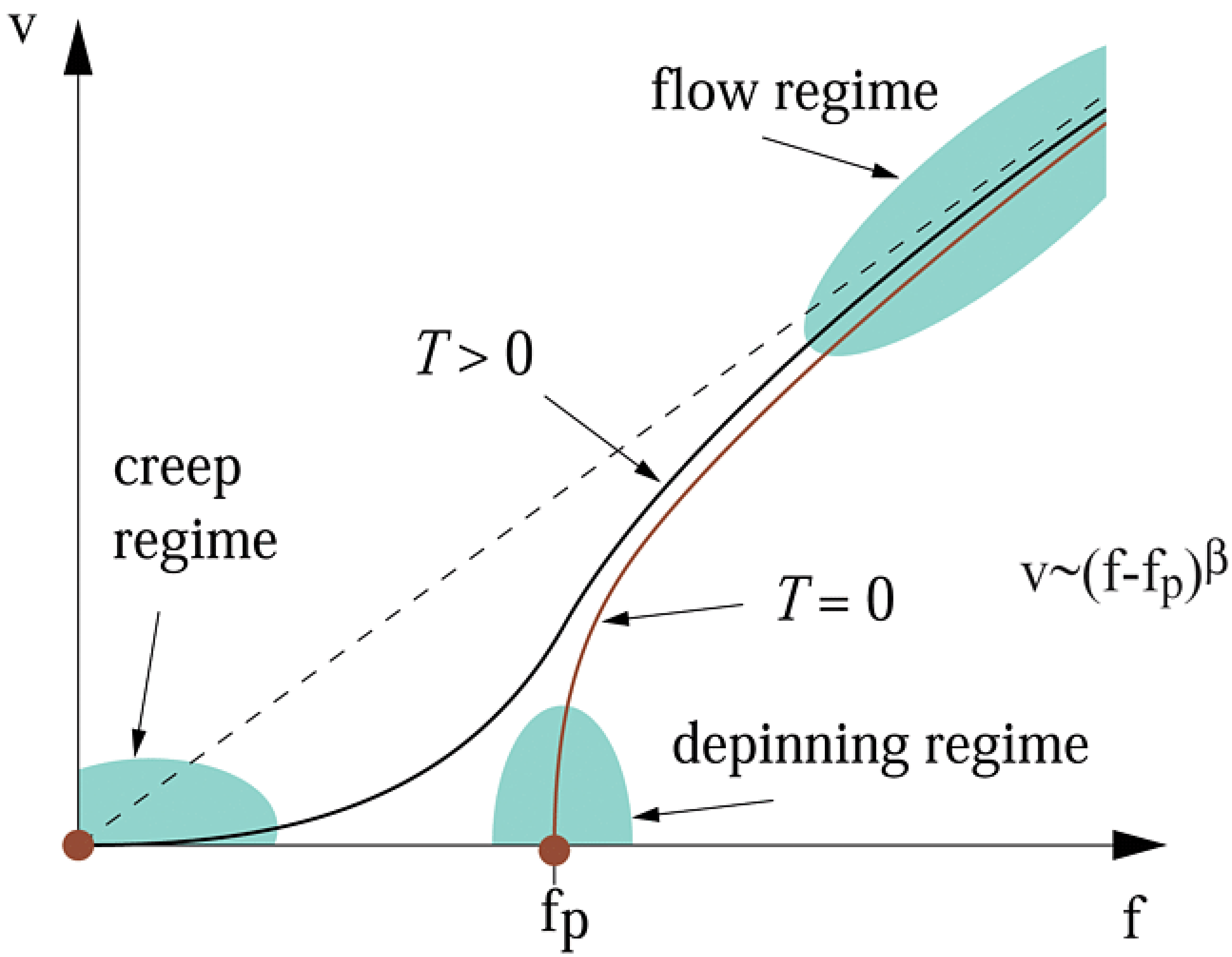


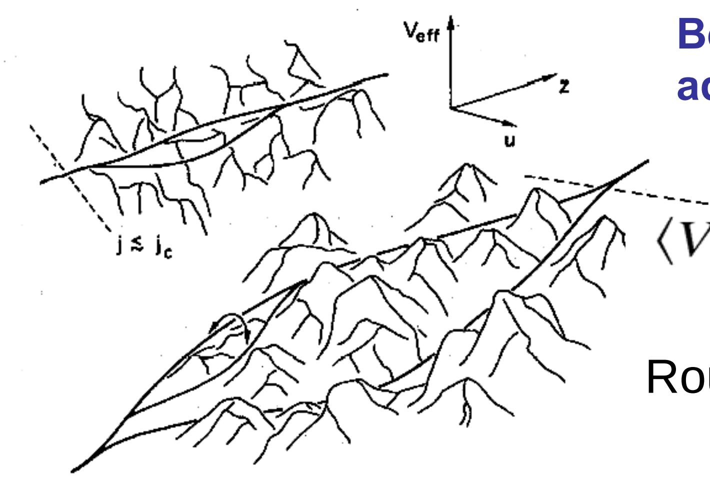
$$w \sim \xi \left(\frac{L}{L_c} \right)^{\varsigma}$$

$\zeta < 1$ is the roughness exponent

$$L_c = (C\xi^2/\Delta)^{2/(4-D)}$$

Dynamic diagram



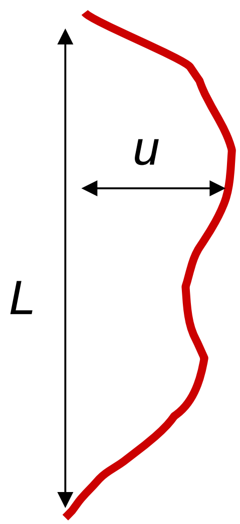


Being placed in disordered medium elastic object adjusts itself to rugged potential relief

$$\langle V(\mathbf{x},\mathbf{u})V(\mathbf{x}',\mathbf{u}')\rangle = \Delta^2\delta^D(\mathbf{x}-\mathbf{x}')f(|\mathbf{u}-\mathbf{u}'|/\xi)$$

Roughness: $w(L) = \langle[\mathbf{u}(\mathbf{x}+\mathbf{L}) - \mathbf{u}(\mathbf{x})]^2\rangle^{1/2}$

On the intermediate scales where $u < a_0$



$$w \sim \xi\left(\frac{L}{L_c}\right)^{\varsigma}$$

$\zeta < 1$ is the roughness exponent

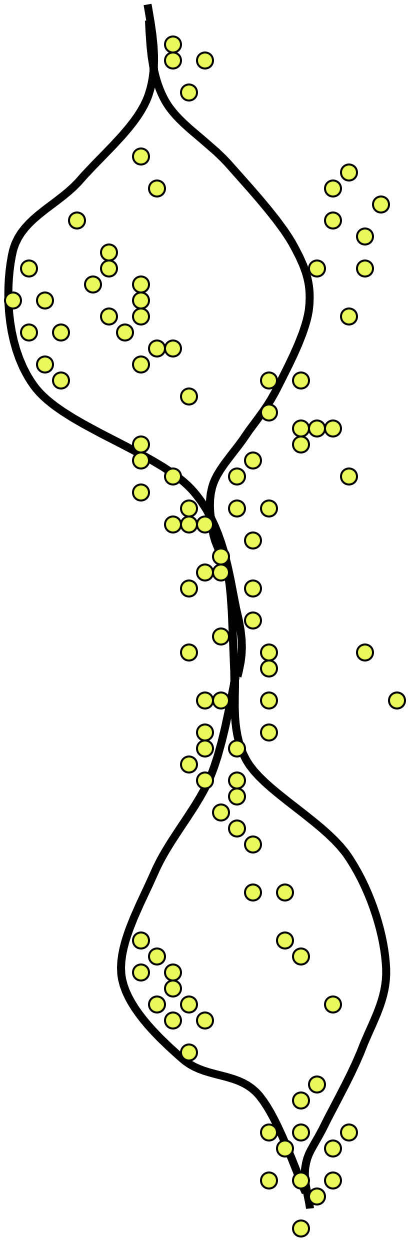
$$E_{barrier} = E_p\left(\frac{L}{L_c}\right)^{\chi} f\left(\frac{x-x'}{w(L)}\right)$$

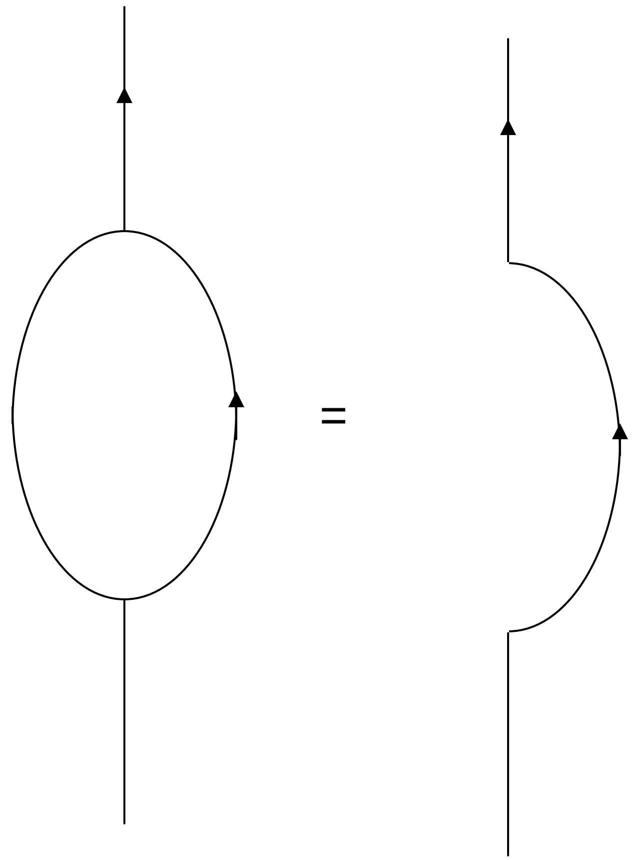
$$L_c = (C\xi^2/\Delta)^{2/(4-D)}$$

$$E_p = a_0\Delta L_c^{D-2}$$

$$\chi = D - 2 + 2\zeta$$

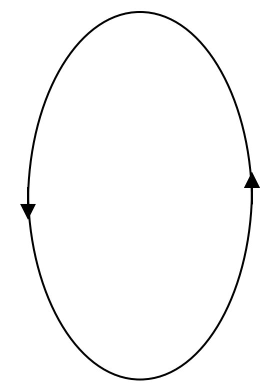
Thermally activated vortex dynamics in random media



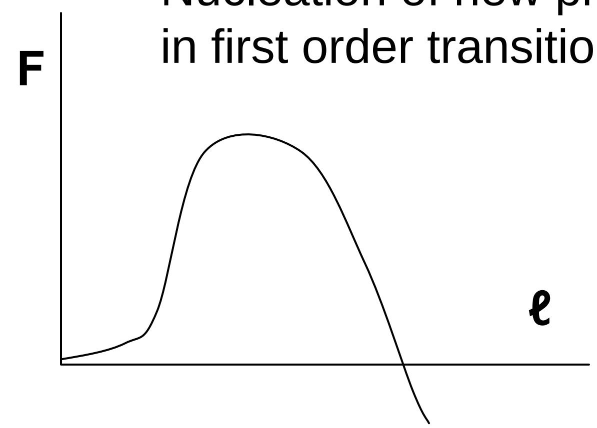


Activated vortex motion:

Creation a vortex loop



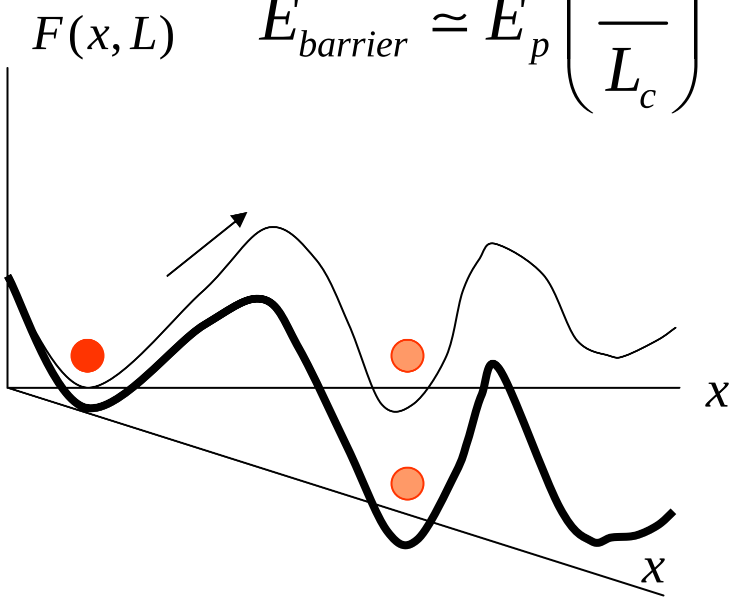
Nucleation of new phase
in first order transitions



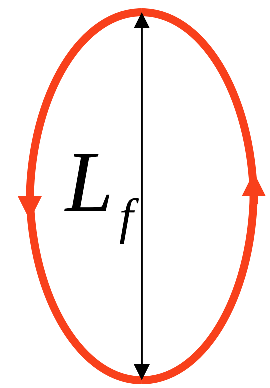
Driving force *f* :

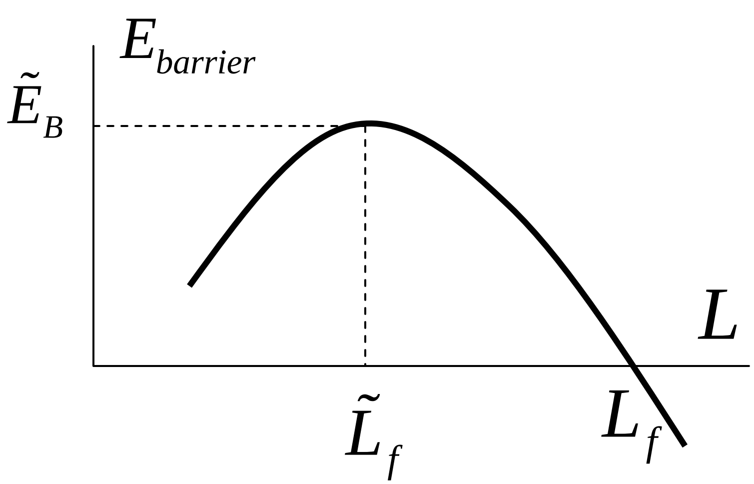
$$E^{(0)}_{barrier} = E_p \left(\frac{L}{L_c} \right)^{\chi}$$

$$E_{barrier} \simeq E_p \left(\frac{L}{L_c} \right)^{\chi} - w(L)\, f L^D = E_p \left(\frac{L}{L_c} \right)^{\chi} \left[1 - \left(\frac{L}{L_f} \right)^{2-\zeta} \right]$$



$$L_f = L_c \left(\frac{f_p}{f} \right)^{1/(2-\varsigma)}$$





Meaning of L_f :

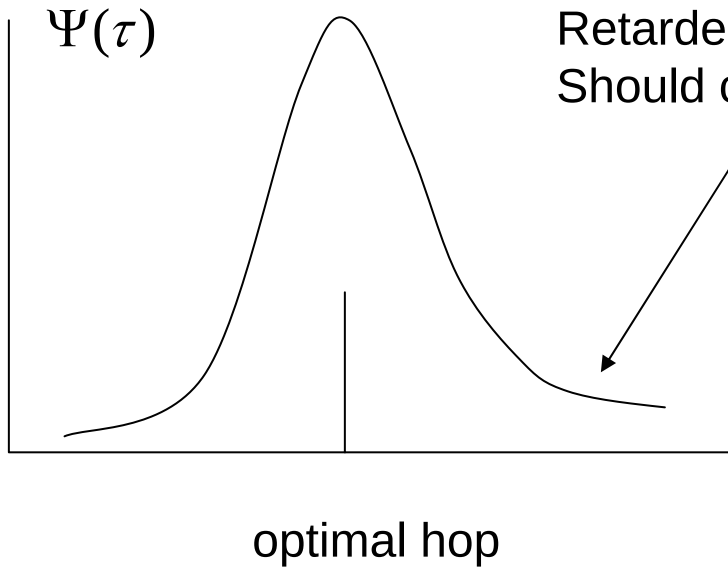
At distances $L > L_f$

pinning is not effective

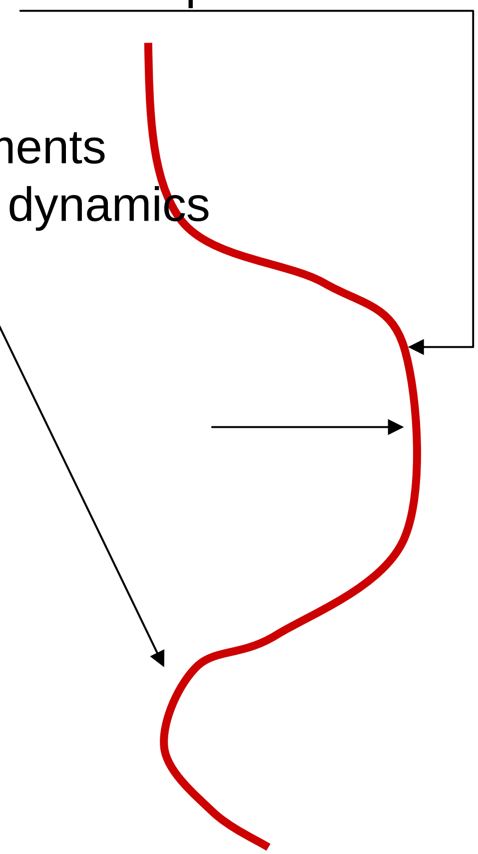
$$E_{barrier}(L_f) \equiv \tilde{E}_B = E_p \left(\frac{f_p}{f} \right)^{\mu}$$

How can it work?

How can our object choose this optimal hop?



Retarded segments
Should control dynamics

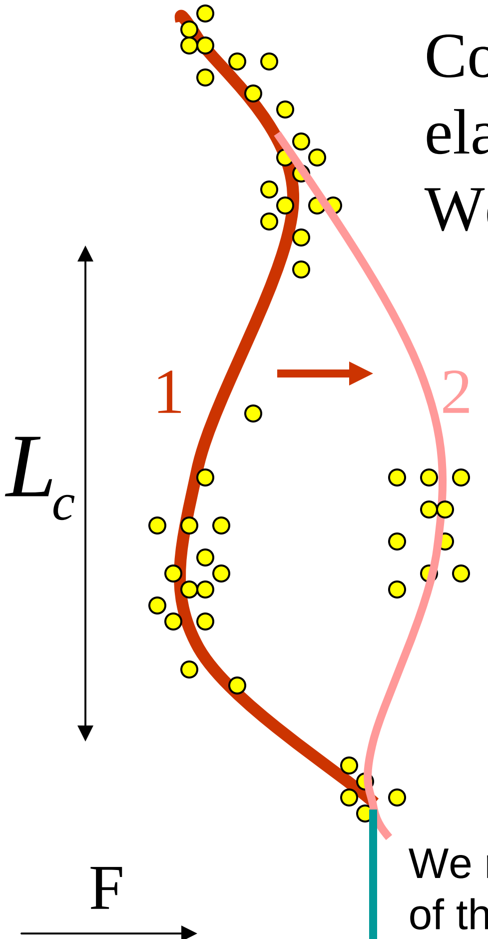


But… for rare fluctuattions $E_{barrier} \sim \log L$

(at best)

Energy gain due to external force $\sim L$

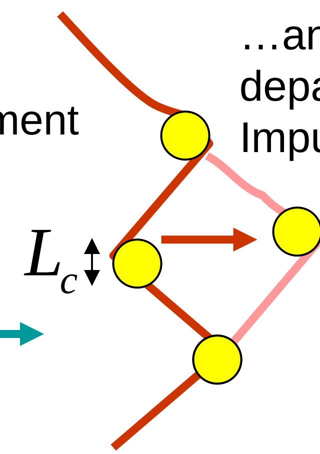
$\Psi(\tau)$?



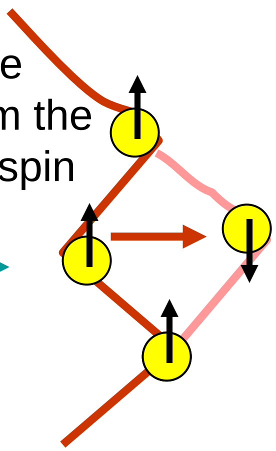
Consider a change in the configuration of the elastic string in the random field:
We call it departure from the initial position 1.

We are interested in the distribution density $\Psi(\tau)$ of the departure times τ. To find it we will determine first the probability distribution $W(E)$ of barriers E controlling string dynamics

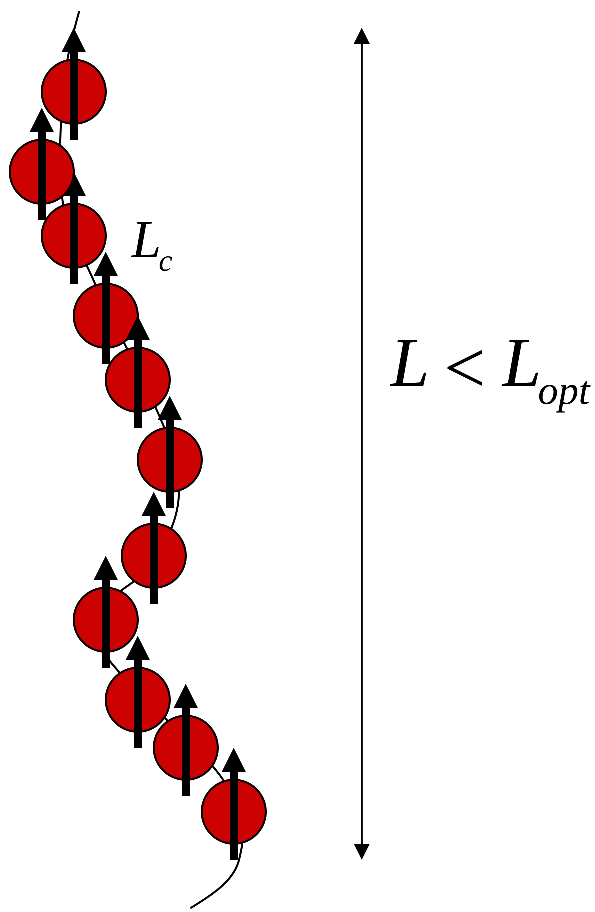
We replace each segment of the length L_c by the strong impurity:

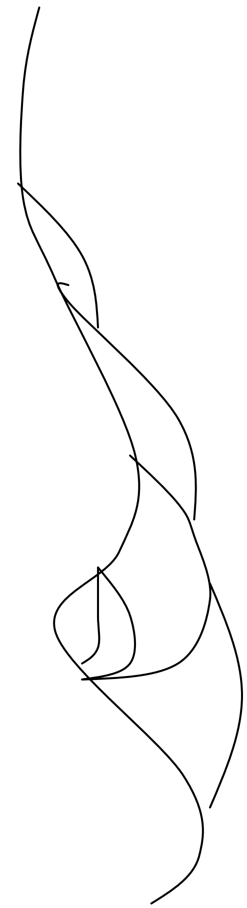


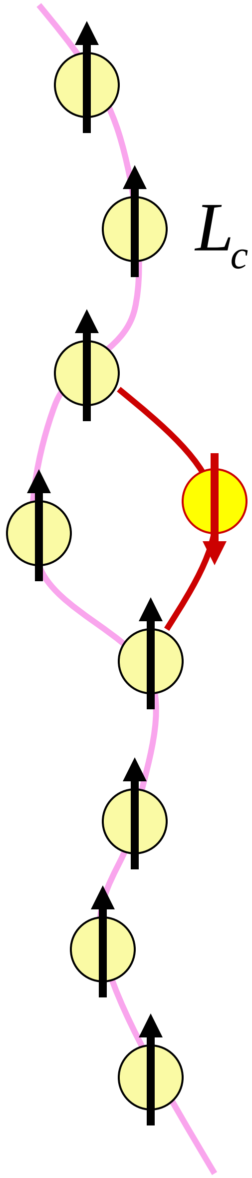
…and view the departure from the Impurity as a spin flip



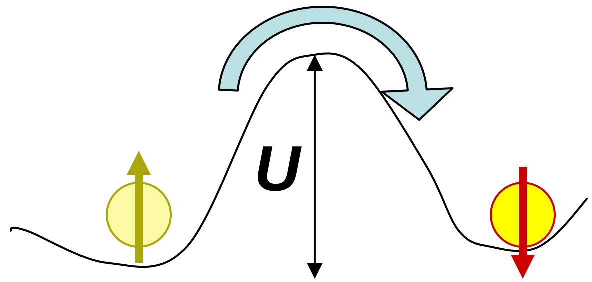
To derive the distribution of energy barriers we define the elementary moving "units" of the string as segments of length L_c. The barriers controlling the hop of these units to the nearest metastable state fluctuate about U_c.





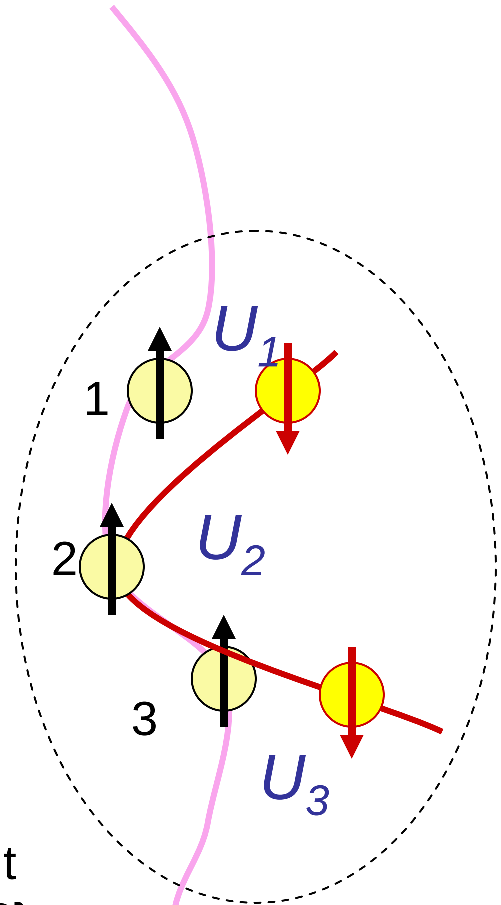


It costs energy U to flip (depart from) a site



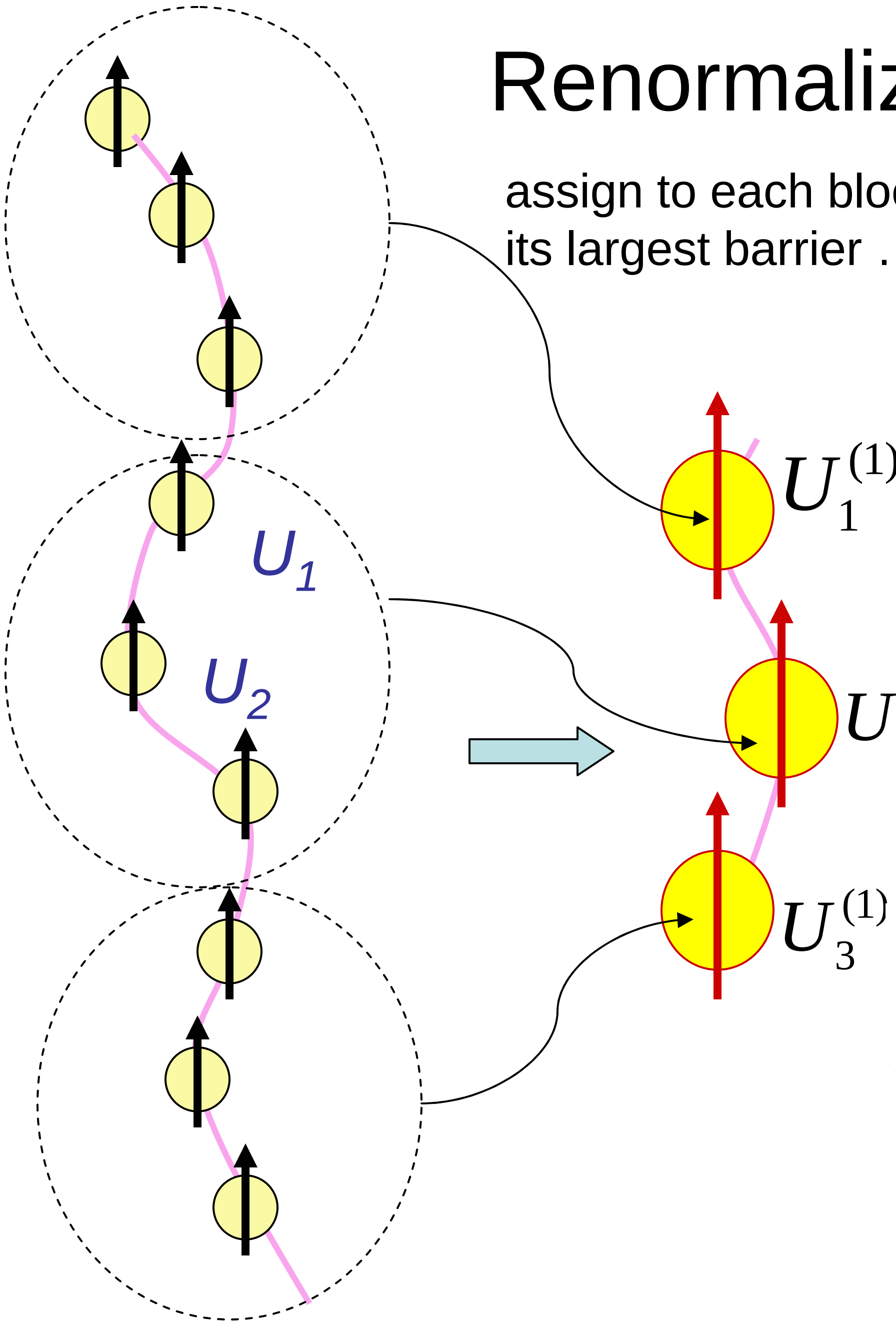
The barriers controlling the flips (hops) to the neighboring site fluctuate about U_c.

Consider a segment of three units: {1,2,3}



Departure of the segment is controlled by $U^{(1)}=max\{U_1,U_2,U_3\}=U_2$

Renormalization procedure



assign to each block
its largest barrier …

replace each block by a new unit
with these new barriers…

…and repeat the procedure…

Generally we consider blocks
consisting of *m* units

By the same argument as above, we find that the hop of a segment L' composed of m blocks L_1 is again controlled by $U' = \max\{U_j^{(1)}\}$. Repeating this procedure, we list all possible configurations of the advancing string L. If for large m a limiting form exists for the distribution function of the energy barriers, this form must be stable under the max operation. In other words if $U = \max\{U_i\}$, then the probability distribution of the extrema U is the same as that of each member U_i of the set.

The probability $\mathcal{P}_L(z)$ that the largest energy barrier controlling the hop of a string of length L is less than U is given by the solution of the functional equation

$$\mathcal{P}_L(z) = \left[\mathcal{P}_L(a_n z + b_n)\right]^n,$$

where $z = (U/U_c - b_n)/a_n$ and all lengths are measured in units of L_c.

The probability distribution of the largest energy barriers for a pinned segment of length L is then,

$$\mathcal{P}_L(U) \sim \exp\left[- e^{-(U - U_c \ln(L/L_c)/U_c} \right] \\ = \exp\left[- (L/L_c) e^{-U/U_c} \right]$$

and the corresponding probability density is given by

$$p_L(U) = \frac{d\mathcal{P}_L}{dU} \sim \frac{L}{L_c} e^{-U/U_c} \exp\left[- (L/L_c) e^{-U/U_c} \right].$$

Global distribution: $$W(U) = \int_{L_c}^{\infty} dL\, n_L \mathcal{P}_L(U)$$

VOLUME 70, NUMBER 5 PHYSICAL REVIEW LETTERS 1 FEBRUARY 1993

Elastic String in a Random Potential

M. Dong, M. C. Marchetti, and A. Alan Middleton
Physics Department, Syracuse University, Syracuse, New York 13244
V. Vinokur
Argonne National Laboratory, Materials Science Division, Argonne, Illinois 60439
(Received 3 January 1992)

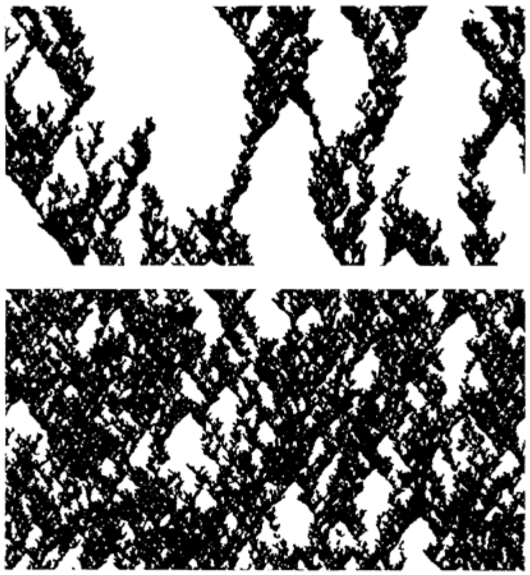


FIG. 2. Maps of string velocity for $f = 0.01$ (top) and $f = 0.076$ (bottom) and $F_p = 0.1$. The vertical axis is time, while the horizontal axis is the position z on the string. Dark regions indicate where the velocity exceeds 0.01. Maps are shown for strings of size $L = 4096$ evolving over a time interval $\Delta t = 30$.

n_L : the density of pinned segments of the length *L*

$$n_L \sim 1/L^{\nu} \qquad \nu = 1 + d/d_f > 2$$

$$\boxed{W(U) \sim e^{-U(\nu-1)/U_c}}$$

$$\tau = \tau_0 \exp(U/T) \quad \Psi(\tau)d\tau = W(U)dU,$$

$$\boxed{\Psi(\tau) \sim T(\tau_0/\tau)^{1+\alpha}, \ \alpha = (\nu - 1)T/U_c}$$

The mean motion is controlled by the largest departure/waiting time corresponding to the hop of the optimal segment $L_{opt}(F)$.

segments on scales $L > L_{\text{opt}}$ slide freely. $\Rightarrow$

$$\langle \tau \rangle \sim \int^{\tau_{\max}} d\tau \Psi(\tau) \tau \sim \exp[(1 - \alpha) U(F)/T]$$

$$\tau_{\max} = \tau_0 \exp[U(F)/T]$$

This is by construction the time over which L_{opt} advances a transverse distance u_{opt}. The mean velocity is then

$$\boxed{v \simeq u_{\text{opt}}/\langle \tau \rangle \simeq \exp[-(1 - \alpha) U(F)/T]}$$

Conclusions:

We have shown that the low-temperature dynamics of driven elastic manifolds in random environment is governed by a power-law distribution of hopping times.

Note: the linear topology of the string is crucial for carrying out the above scaling procedure. For a D-dimensional manifold the network of metastable states generally forms a multi-connected cluster. Following H.J.Hermann and H.E.Stanley, PRL 53, 1121 (1984) one can however describe this cluster as a network of nodes connected by one-dimensional links. Each link consists of a sequence of multiply connected subclusters (beads) and singly connected sections. The backbone connecting two randomly chosen sites can be then described as singly connected upon rescaling. Thus we recover our result.

We therefore expect that the exponential distribution for controlling barrier is generic for all random correlated systems.