Nonlinear excitations as tools to analyze DNA thermodynamics and dynamics.

Michel Peyrard
Ecole Normale Supérieure de Lyon, France

in collaboration with

Thierry Dauxois and Santiago Cuesta Lopez
Ecole Normale Supérieure de Lyon, France

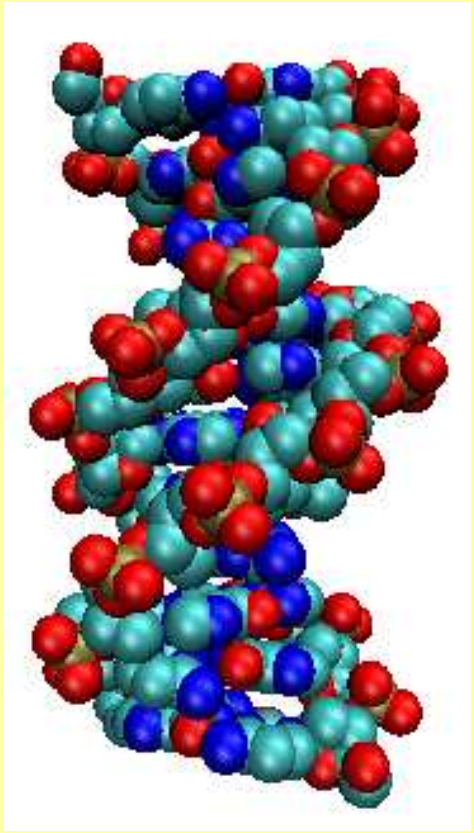
Nikos Theodorakopoulos
Theoretical and Physical Chemistry Institute, Athens, Greece

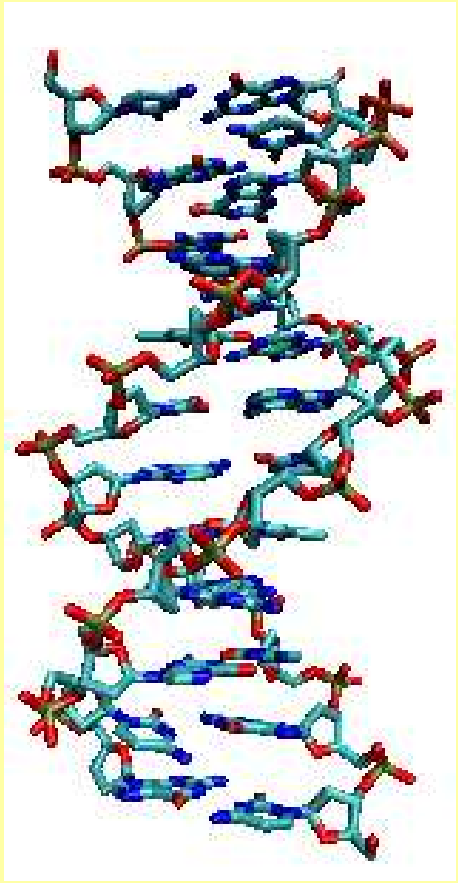
Guillaume James
Institut National Polytechnique de Grenoble and CNRS, France

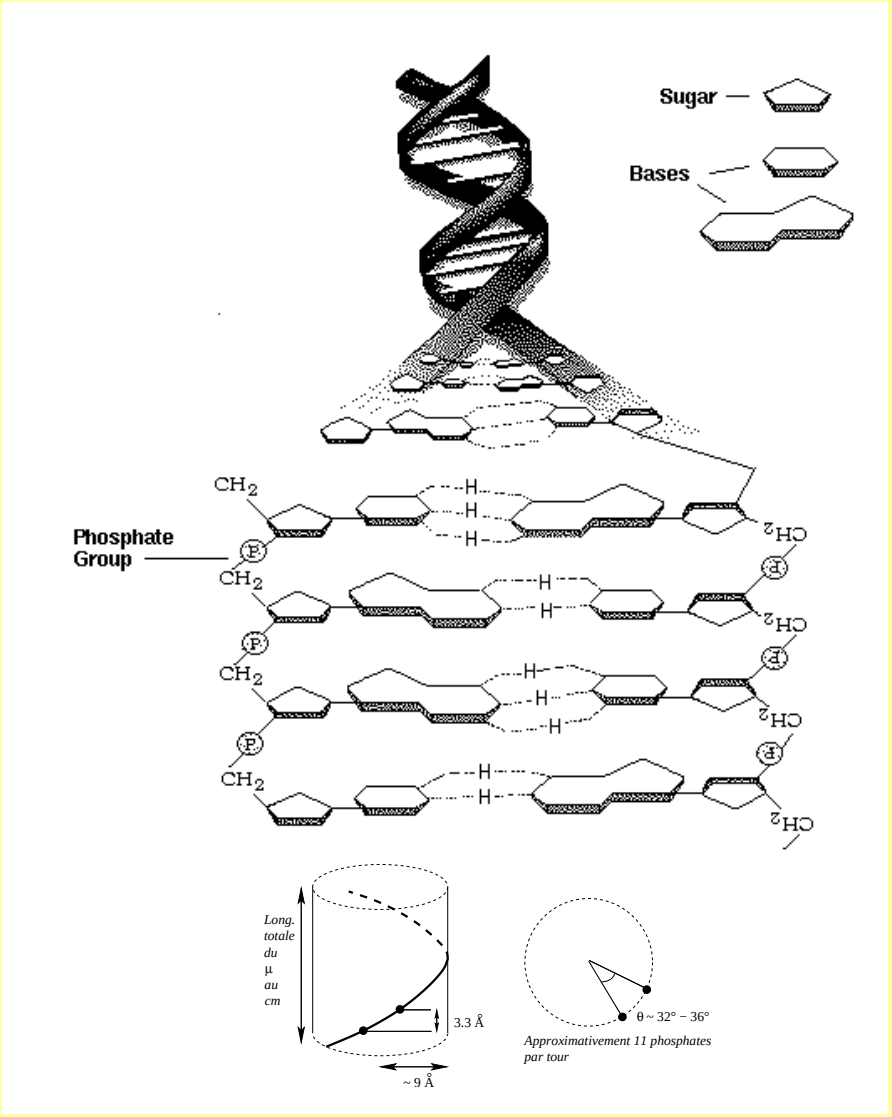
To Pasquale, colleague and friend!

- **Questions raised by DNA properties**
  - some experimental facts
- **A nonlinear model for DNA fluctuations and melting.**
  - Nonlinear localization
  - DNA melting, a phase transition in one dimension
- **From statistics to dynamics: limitations of the model.**
  - An improved model and its dynamics
  - A new class of discrete breathers
- **Sequence effects in DNA: another challenge.**

Questions raised by DNA properties





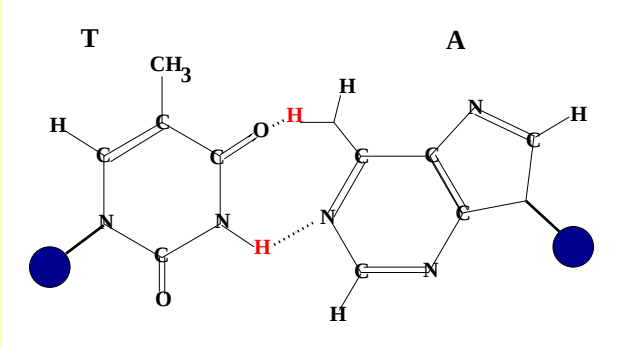


DNA is a highly dynamical entity.

“Breathing of DNA” has been known from biologists for decades.

- Observation: kinetic of imino-proton exchanges

Guéron et al. used NMR.
(Gueron et al. Nature **328**, 89 (1987))

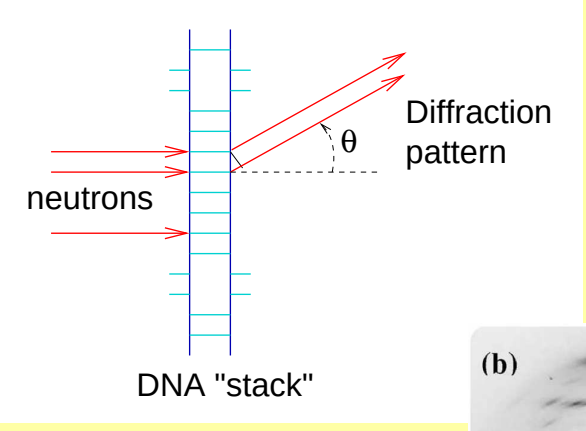


- Lifetime of a base-pair (closed time): 10 ms.
- **Individual** base pair opening (neighbors have different lifetimes)

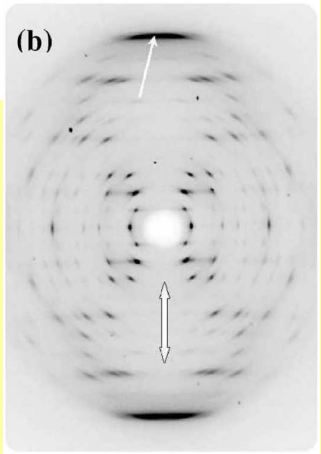
→ ***DNA as a “model lattice” for nonlinear science***

DNA fluctuations can be studied with various experimental methods

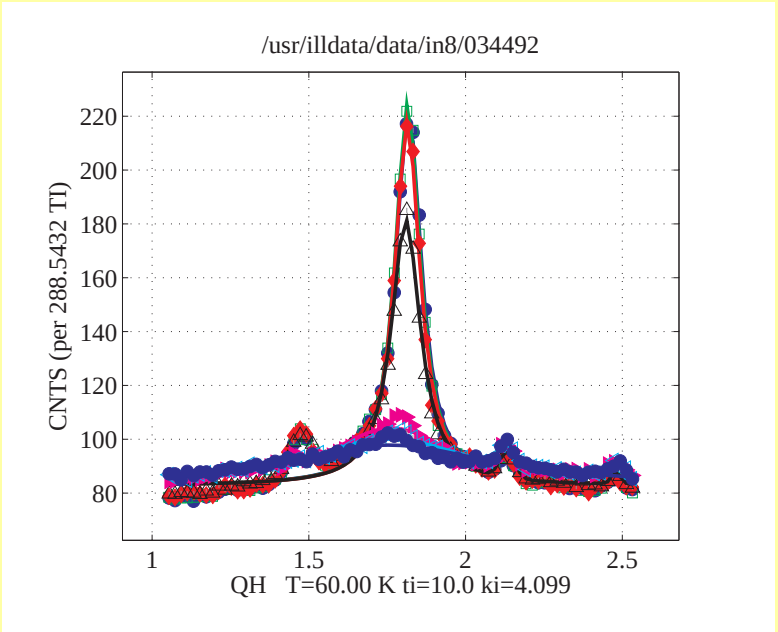
Neutron diffraction (with Andrew WILDES, ILL, Grenoble)



Measure the coherence length of the base stacking, which makes a diffraction grating.



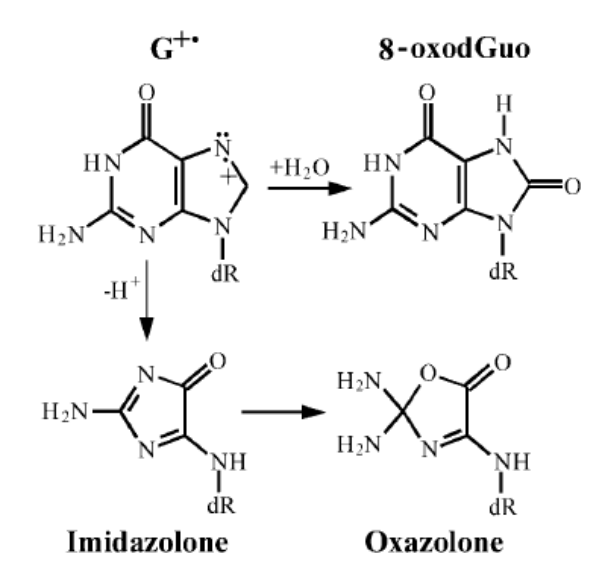
M. Krisch, et al. Phys. Rev. E **73** 061909-1-10 (2006)



$T = 60, 75, 78, 79, 80, 81\ ^{\circ}\mathrm{C}$

UV Laser oxidative Guanine modifications: a probe of local stacking fluctuations. (with Dimitar ANGELOV, ENS Lyon)

High-intensity UV laser pulses → guanine oxidative lesions strongly modulated by the local structure.



Oxozalone sensible to cleavage by piperidine glycosylase

8-oxodG sensible to cleavage by Fpg protein, and appears only in helicoidal stacking

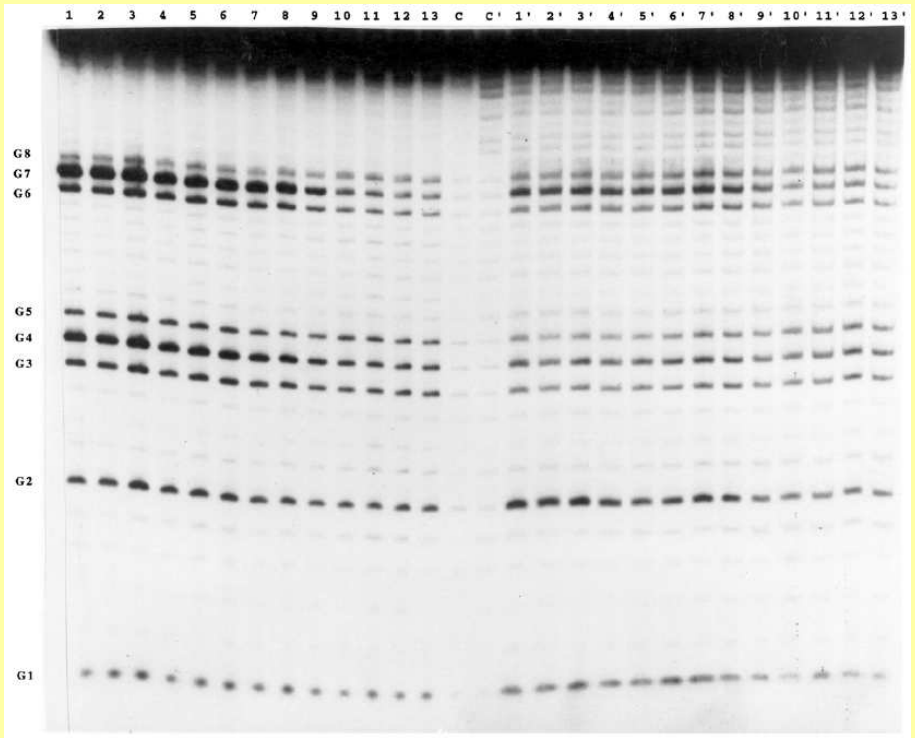
→ R_{pip}/R_{FpG} indicative of local opening

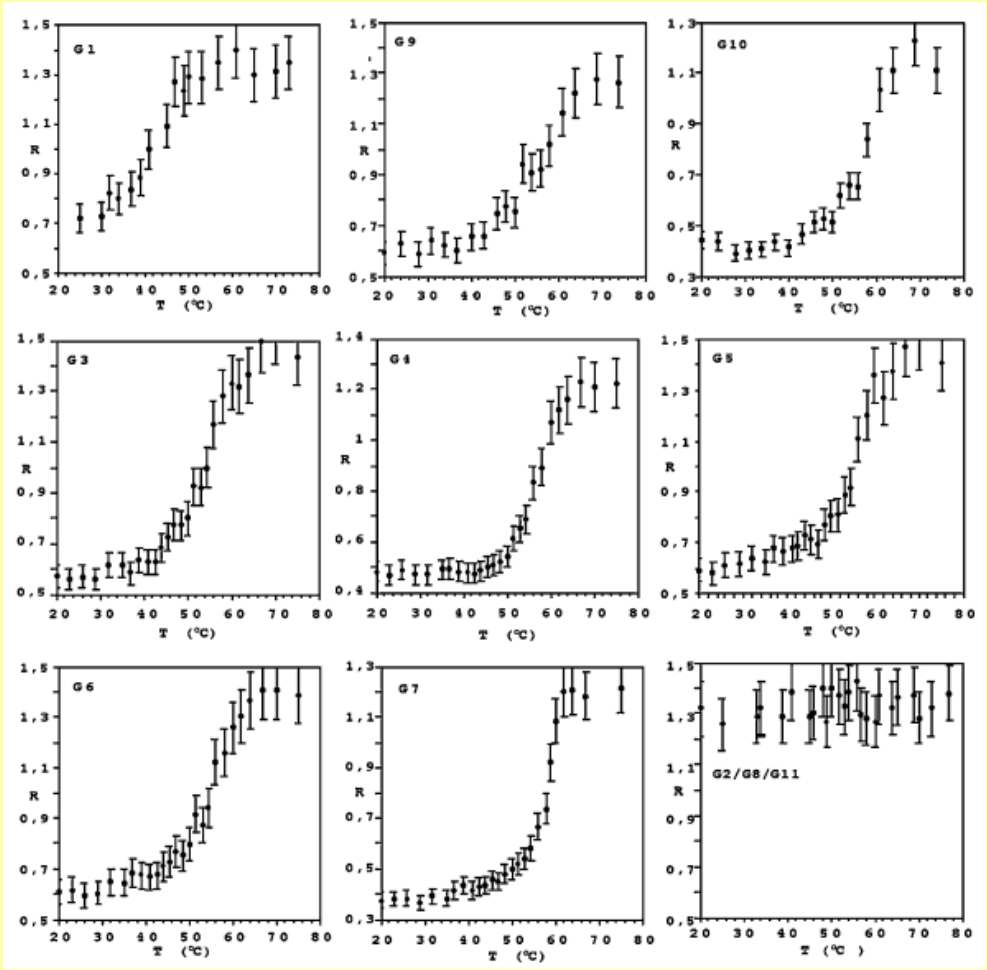
Opens the possibility to evaluate fluctuational opening versus sequence.

From A. Spassky and D. Angelov, J. Mol. Biol. **323** 9 (2002)

```
                1     2    345         678
5'-ATAAAAGTTACGCTAGGGATAACAGGGTAATATAACG-3'
3'-TATTTTCAATGCGATCCCTATTGTCCCATTATATTGC-5'
                  11 10          9
```

FpG protein hot piperidine



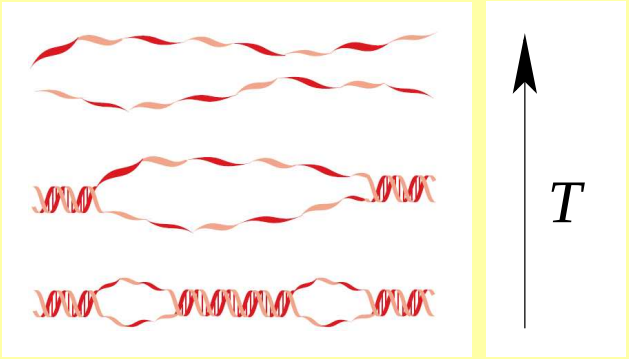


Provides data for all *GC* pairs, contrary to fluorophore–quencher methods.

DNA fluctuations are important for biological function

The genetic code is buried (protected) inside the structure

Local opening can be thermally created (observed in "DNA melting")
(M. Peyrard, Nature Physics, **2**, 13-14 (2006))



- Also observed in homopolymers (not disorder-induced localization)
- Larger openings: DNA "bubbles"

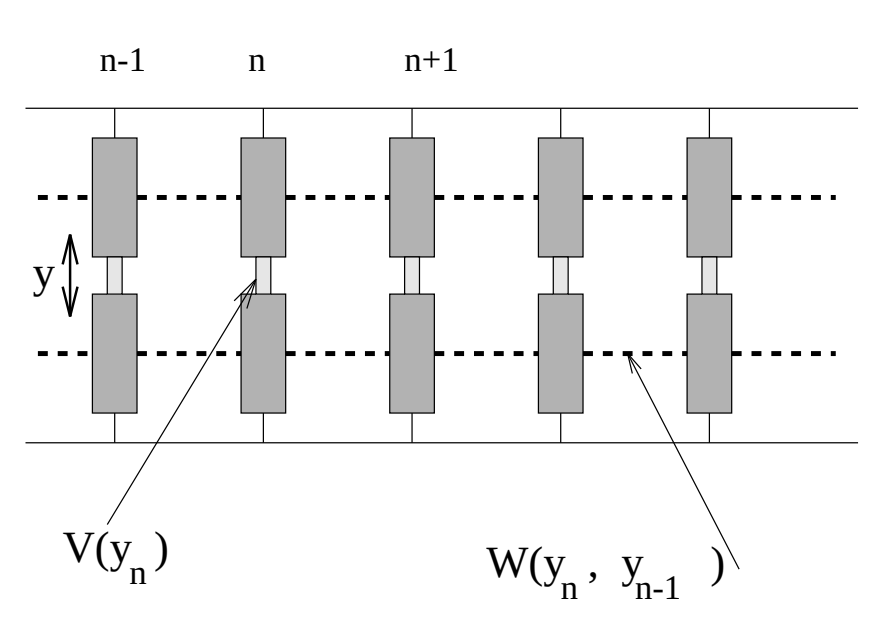
Questions

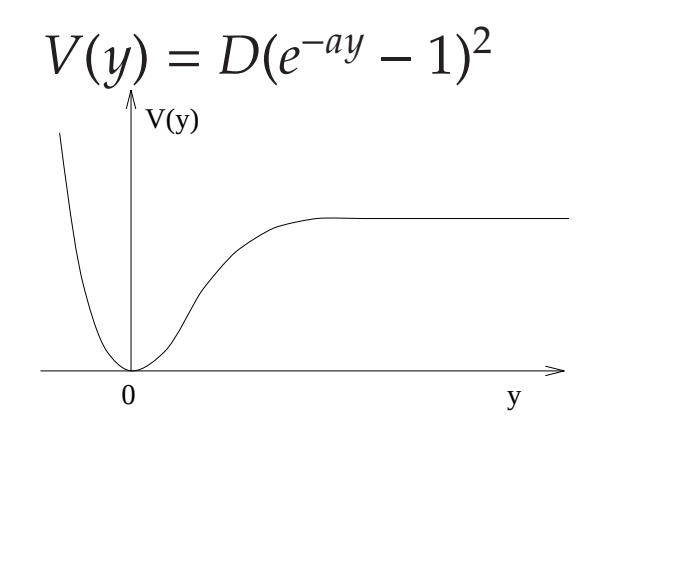
- How can we understand the localization observed in DNA melting?
- Can we understand the dynamics of the fluctuations (lifetime of a closed base pair, and of the open state)

A simple nonlinear model for DNA fluctuations and melting

(M. Peyrard and A.R. Bishop, Phys. Rev. Lett. **62**, 2755 (1989))

Beyond Ising: describe dynamics → real variable y for base-pair stretching

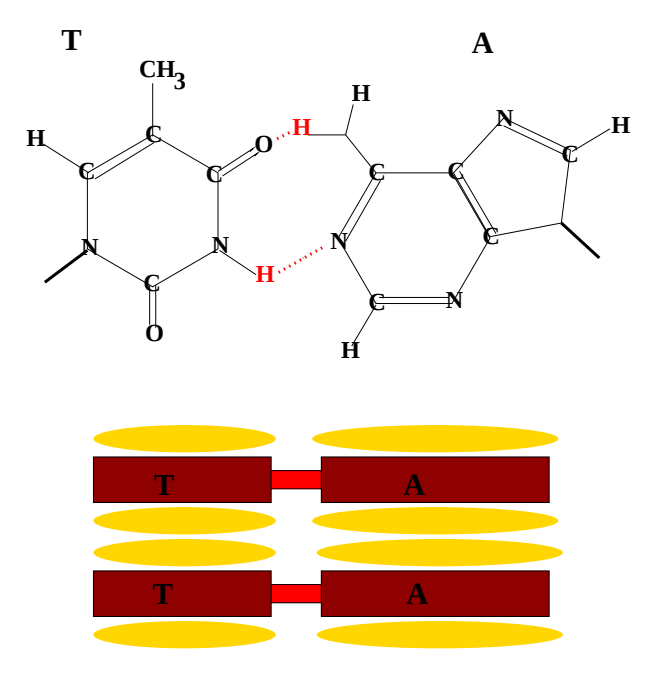




DNA structure only encoded in the potentials

$$H = \sum_n \left[\frac{1}{2} m \left(\frac{dy_n}{dt} \right)^2 + W(y_n, y_{n-1}) + V(y_n) \right]$$

Coupling potential W (stacking) : crucial for the statistical mechanics



Type I:

$$W(y_n, y_{n-1}) = \frac{1}{2} K \, (y_n - y_{n-1})^2$$

Type II:

$$W(y_n, y_{n-1}) = \frac{1}{2} K \left[1 + \rho e^{-\alpha (y_n + y_{n-1})} \right] (y_n - y_{n-1})^2$$

if y_n **and** $y_{n-1} \ll 1/\alpha \Rightarrow K' \approx K(1+\rho)$
if y_n **or** $y_{n-1} \gg 1/\alpha \Rightarrow K' \approx K$

Model parameters:

Actual potential interactions → can be *estimated*
(ex. D energy for base-pair breaking ≈ 0.03 eV. Includes solvent effects)

Comparison with experiments → precise their values

Dynamics of the thermalized model (homopolymer)

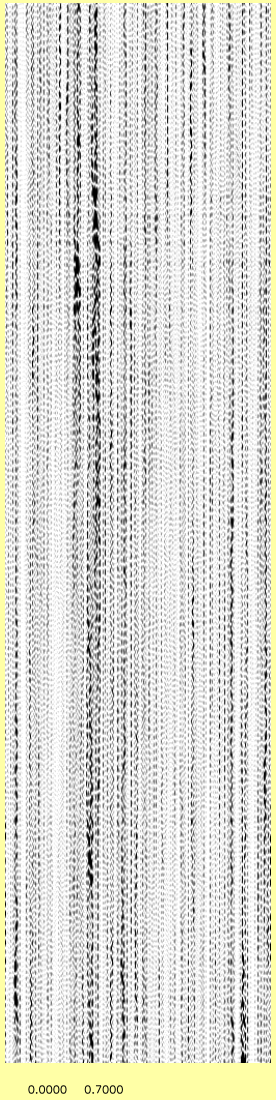
Simulation at constrained T

- Low temperature:
  Thermally generated discrete breathers
  corresponds to "DNA breathing" of biologists.

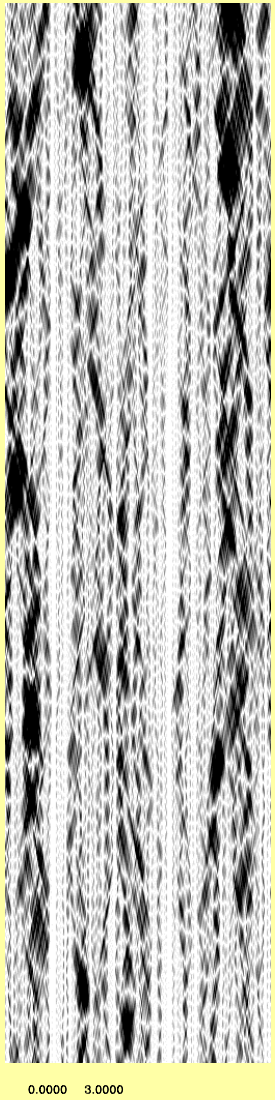
- High temperature: "patches" of high and low "T"
  corresponds to "DNA denaturation bubbles"

- The thermal denaturation: a true transition in 1D?
  (not only an academic problem!)

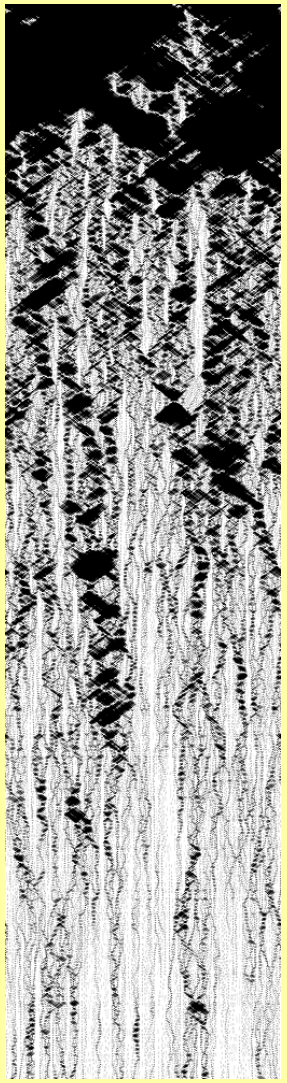
T = 100 K



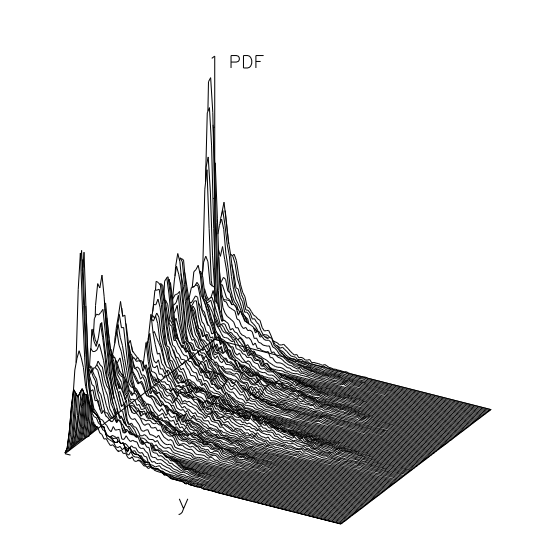
T = 340 K



increasing T



The study of the biological problem leads to a new idea: temporary localization of thermal fluctuations.

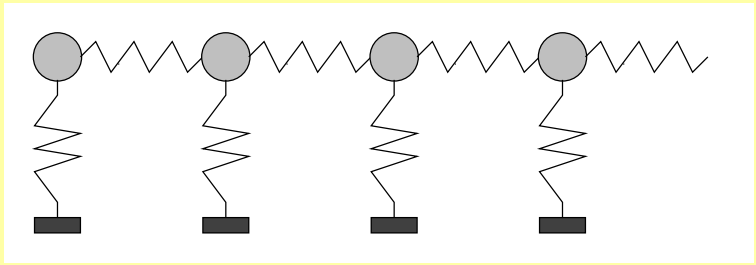


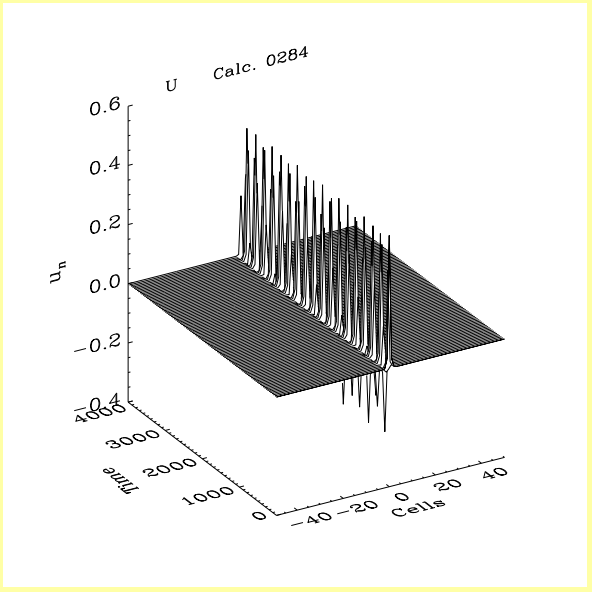
Energy equipartition temporarily broken!

Local distribution $\mathcal{P}(y, n_0, t_0, \Delta n, \Delta t)$

Behind this phenomenon: discrete breathers:

$$m\ddot{u}_n - K(u_{n+1} + u_{n-1} - 2u_n) + \omega_d^2(u_n - u_n^3) = 0$$



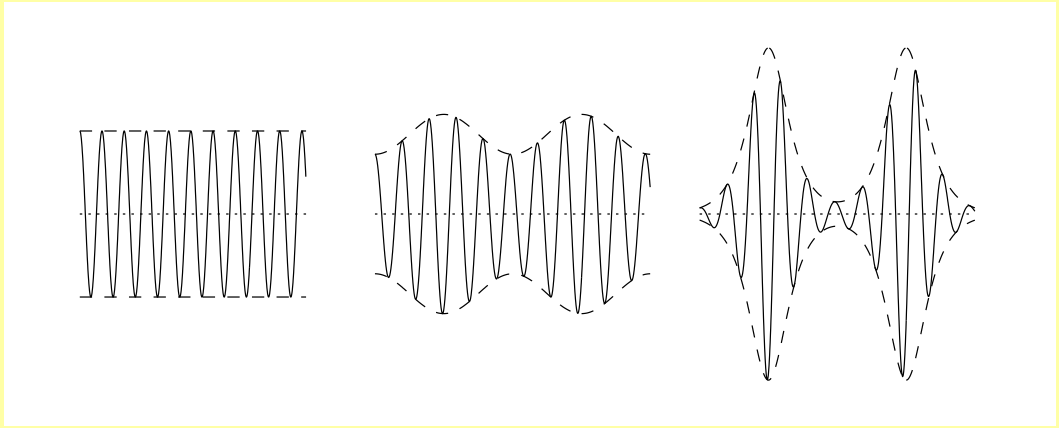


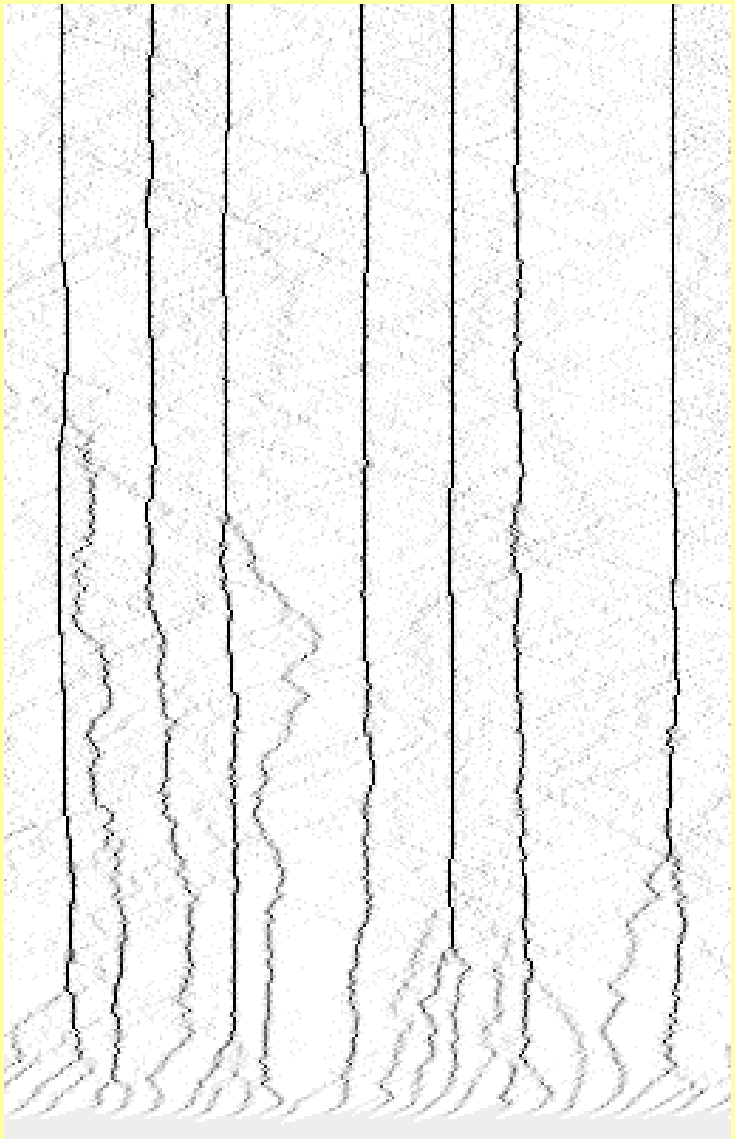
Very general, no strict condition contrary to solitons

How does it occur ? Nonlinear localization.

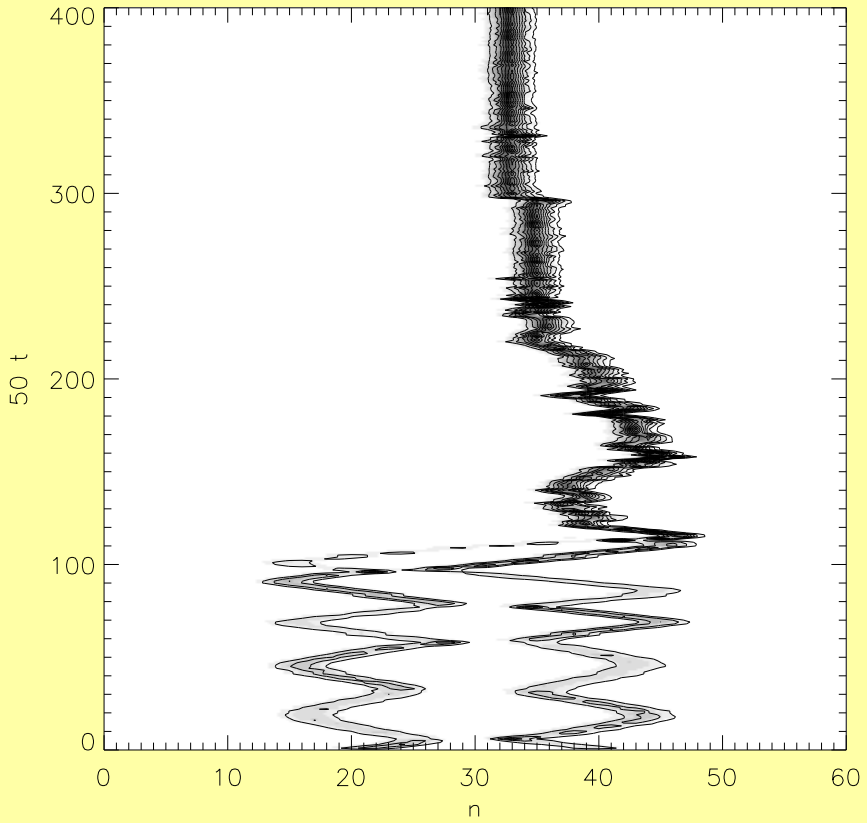
Time evolution of the energy density shows the formation of local modes through two mechanisms:

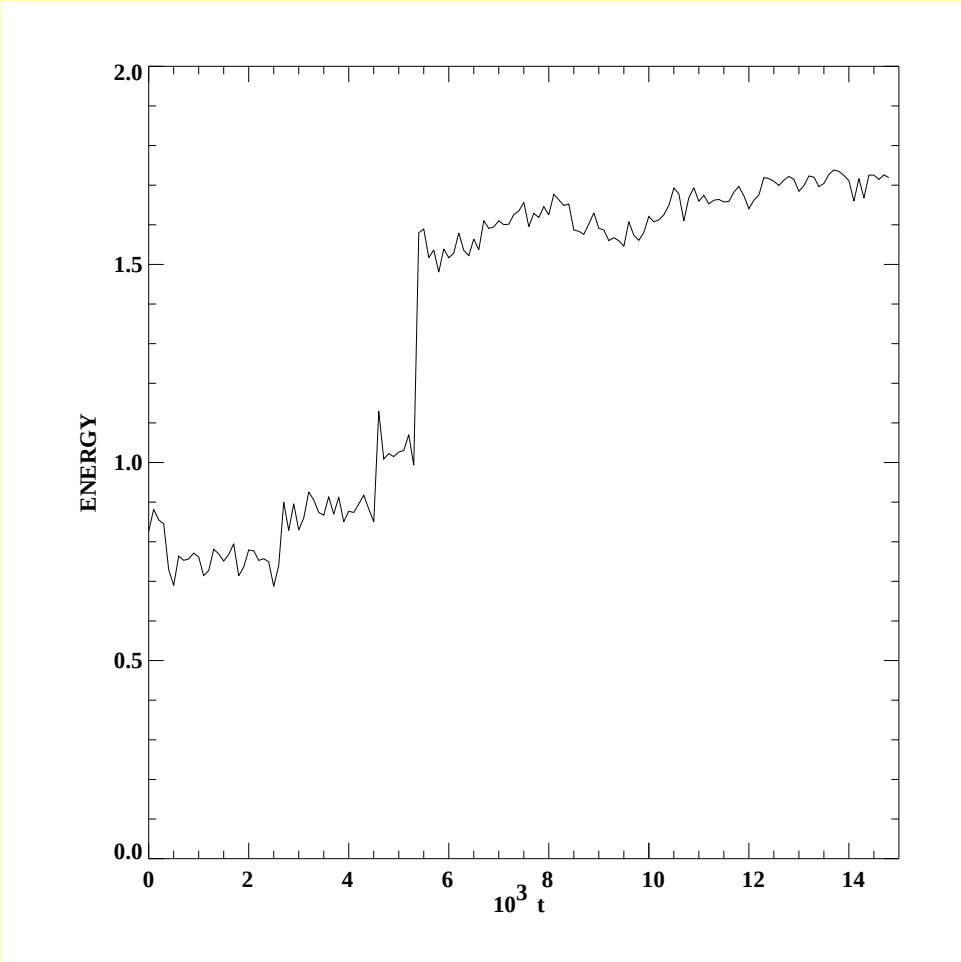
- Modulational instability of a plane wave.
- Energy exchange in collisions.





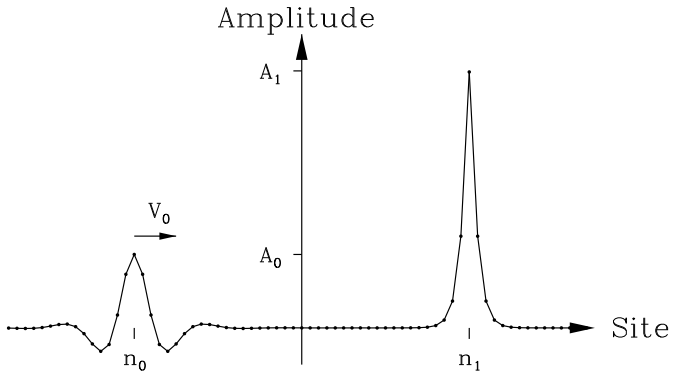
Discreteness helps **collecting** energy: discrete breathers are **NOT** solitons.



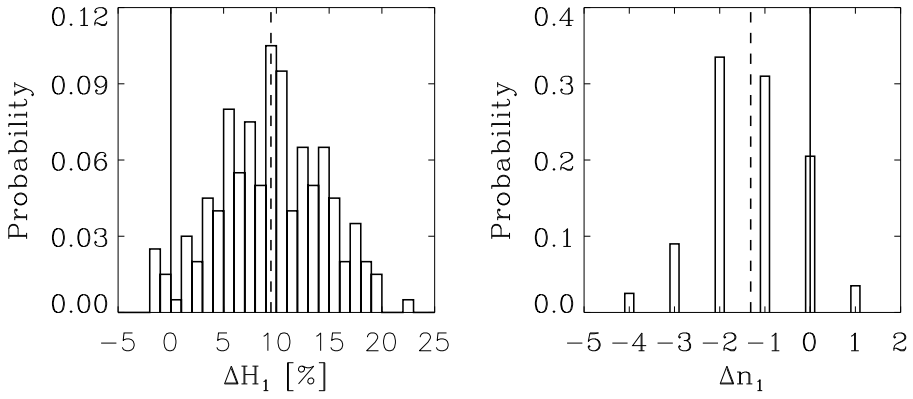


The biggest (most stable) is favored.

Statistics of the collisions (200 collisions)

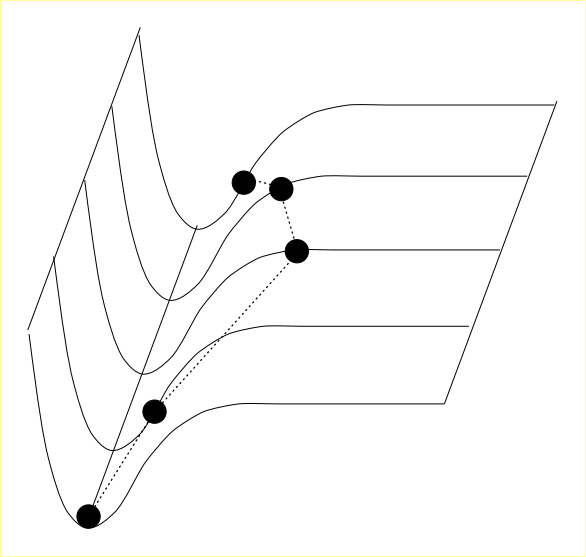


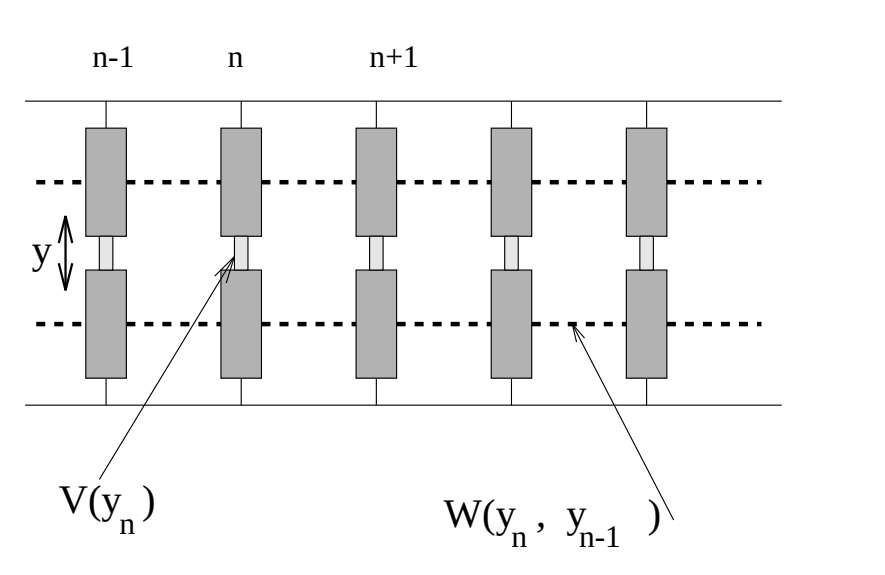
Energy exchange



$$\Delta H_1 = \frac{E_{\text{big}}(\text{after}) - E_{\text{big}}(\text{before})}{E_{\text{small}}(\text{before})}$$

The statistical physics of DNA “melting”





Compute the partition function $\mathcal{Z}$ and then the free energy $F = -k_B T \ln \mathcal{Z}$

$$\mathcal{Z} = \int_{-\infty}^{+\infty} \prod_{n=1}^{N} dy_n dp_n e^{-\beta H} = (2\pi m k_B T)^{N/2} \times \int_{-\infty}^{+\infty} \prod_{n=1}^{N} dy_n e^{-\beta[W(y_n, y_{n-1}) + V(y_n)]} \delta(y_N - y_1)$$

Method: Transfer integral calculation:

$$\mathcal{Z} = (2\pi m k_B T)^{N/2} \mathcal{Z}_y = (2\pi m k_B T)^{N/2} \int_{-\infty}^{+\infty} \prod_{n=1}^{N} dy_n e^{-\beta[W(y_n, y_{n-1}) + V(y_n)]} \delta(y_N - y_1)$$

Define the transfer operator $y_{n-1} \to y_n$ and its eigenfunctions ϕ_i by:

$$\int dy_{n-1} \exp -\beta\Big\{W(y_n, y_{n-1}) + V(y_n))\Big\} \times \phi_i(y_{n-1}) = e^{-\beta\varepsilon_i}\phi_i(y_n)$$

Expand $\delta(y_N - y_1) = \sum_i \phi_i^*(y_N)\phi_i(y_1)$ and perform successively the integrals over $y_1, y_2, \ldots, y_{N-1}$

$$\mathcal{Z}_y = \sum_i e^{-\beta N \varepsilon_i} \int dy_N |\phi_i(y_N)|^2 = \sum_i e^{-\beta N \varepsilon_i} \qquad \text{because } \phi_i \text{ normalized.}$$

In the thermodynamic limit ($N \to \infty$) $\mathcal{Z}_y = e^{-\beta N \varepsilon_0}$ (ε_0 smallest eigenvalue of the TI operator)

Free energy per site $f = -\dfrac{k_B T}{N} \mathcal{Z} = \varepsilon_0 - \dfrac{k_B T}{2} \ln\left(2\pi m k_B T\right)$

The second term is non-singular with T. Therefore, if there is a phase transition, it shows up in the variation of ε_0 versus T.

The problem is thus to find the eigenvalues (and eigenfunctions) of the transfer integral operator.

Similar calculations give

- The order parameter $\sigma = \langle y \rangle = \int_{-\infty}^{+\infty} dy\ y\ |\phi_0|^2$
- Correlation functions

$$C(n) = \langle (y_n - \sigma)(y_0 - \sigma) \rangle = \sum_{i=1}^{+\infty} |\langle \phi_i | y | \phi_0 \rangle|^2 e^{-\beta(\varepsilon_i - \varepsilon_0)|n|}$$

Solving for the eigenstates of the TI operator:
For the harmonic coupling $W(y_n, y_{n-1}) = \frac{1}{2}K(y_n - y_{n-1})^2$ the calculation can be made analytically in the limit of strong coupling between sites (bad for DNA, but useful to understand qualitatively).
We want to solve

$$\int_{-\infty}^{+\infty} dx\ e^{-\beta\left[\frac{1}{2}K(y-x)^2 + V(y)\right]} \phi(x) = e^{-\beta\varepsilon}\phi(y)$$

Where x is base-pair stretching at site $n-1$ and y the stretching at site n. If K is large enough we can define z by $x = y + z$ and expand $\phi(y+z)$ in powers of z

$$\int_{-\infty}^{+\infty} dz\ e^{-\beta\left[\frac{1}{2}Kz^2\right]} \left[\phi(y) + z\phi'(y) + \frac{1}{2}z^2\phi''(y)\right] = e^{-\beta[\varepsilon - V(y)]}\phi(y)\ .$$

Then we perform the Gaussian integrals in z (the odd one vanishes)

$$\sqrt{\frac{2\pi}{\beta K}}\left[\phi(y) + \frac{1}{2\beta K}\phi''(y)\right] = e^{-\beta[\varepsilon - V(y)]}\phi(y)$$

$$\sqrt{\frac{2\pi}{\beta K}}\left[\phi(y)+\frac{1}{2\beta K}\phi''(y)\right]=e^{-\beta[\varepsilon-V(y)]}\phi(y)$$

The prefactor can be integrated in the exponent of the r.h.s. as

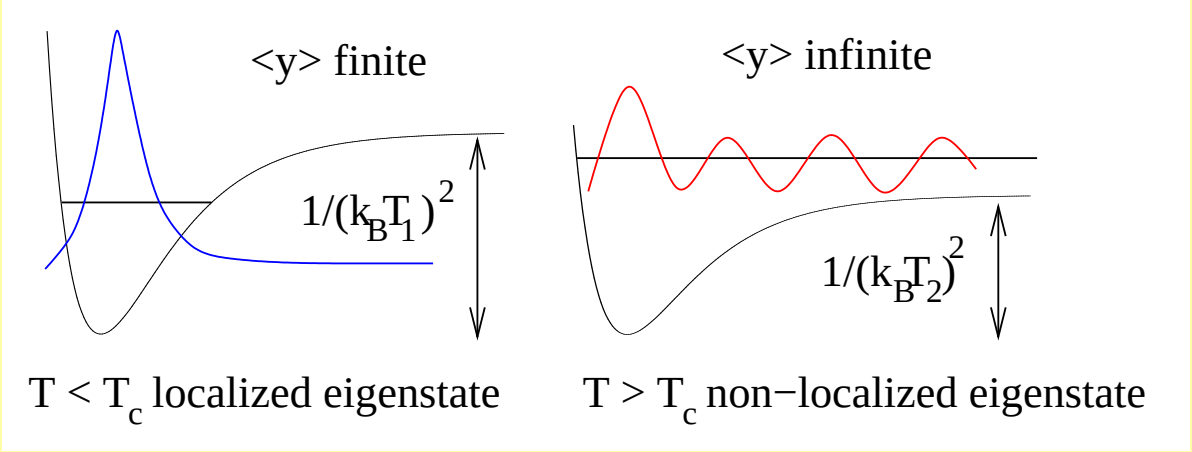
$$\left[\phi(y)+\frac{1}{2\beta K}\phi''(y)\right]=\exp\left\{-\beta\left[\varepsilon+\frac{1}{2\beta}\ln(2\pi/\beta K)-V(y)\right]\right\}\phi(y)$$

Define $\tilde{\varepsilon}=\varepsilon+\frac{1}{2\beta}\ln(2\pi/\beta K)$ and expand the exponential if $\beta D<1$ (the Morse potential is bounded by D)

$$\left[\phi(y)+\frac{1}{2\beta K}\phi''(y)\right]\approx\left[1+\beta V(y)-\beta\tilde{\varepsilon}\right]\phi(y)$$

We then obtain an equation formally identical to the Schrödinger equation of a particle in the potential $V(y)$:

$$\boxed{-\frac{1}{2K}\phi''(y)+\beta^2V(y)\phi(y)=\beta^2\tilde{\varepsilon}\phi(y)}$$

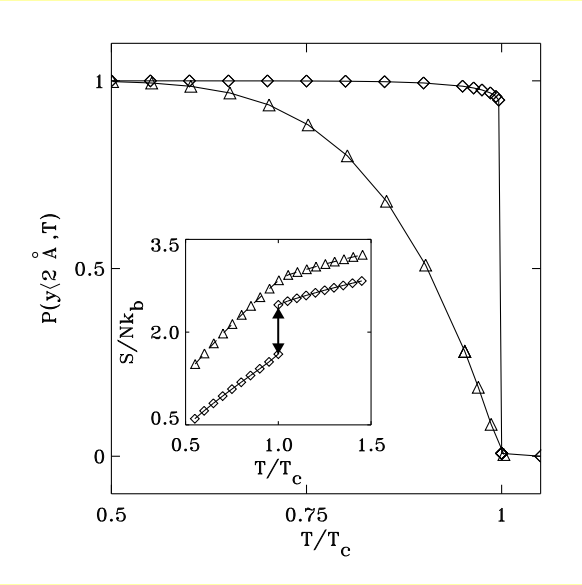


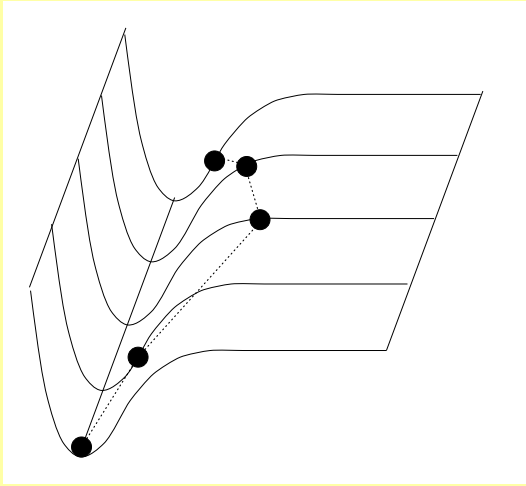
At low T the effective Morse potential is deep enough to have a localized ground state.

For $T>T_c=\frac{2\sqrt{2KD}}{ak_B}$ the ground state is non-localized $\Rightarrow$ denaturation

Discussion of the results.

The transition becomes very sharp when we introduce anharmonic stacking: agrees with experiments.





$$F = U - TS$$

Application: the model can be used to model actual DNA denaturation experiments (heteropolymers)

In this case the transfer integral is calculated numerically exactly

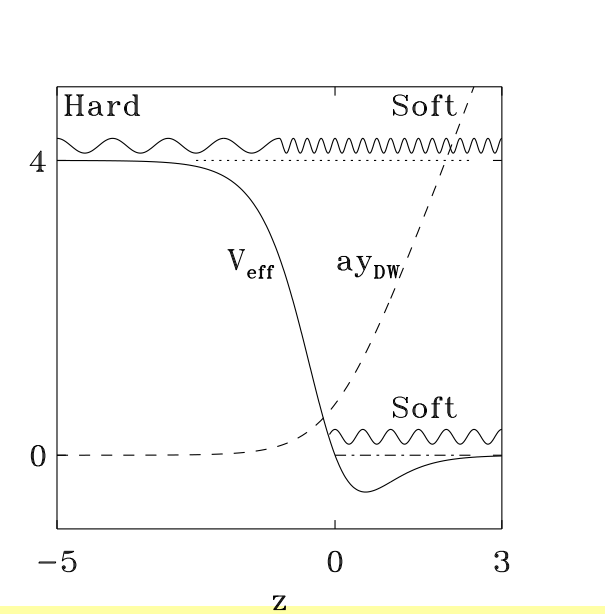
T.S. van Erp, S. Cuesta-Lopez, J.-G. Hagmann, M. Peyrard, Phys. Rev. Lett. **95**, 218104 (2005))

Another point of view: nonlinear science helping statistical mechanics.

(T. Dauxois, N. Theodorakopoulos, M. Peyrard, J. Stat. Phys. **107** 869 (2002))

Transition due to the motion of *Domain Walls*

$$y_{DW}^{\pm}(x) = \frac{1}{a}\ln\left[1 + e^{\sqrt{S/2}\,(x-x_0)}\right] \qquad \text{with} \qquad S = \frac{K}{Da^2}$$



$$E_{DW}^{+}(x_0 \pm \ell) - E_{DW}^{+}(x_0) = \mp 2\ell D$$

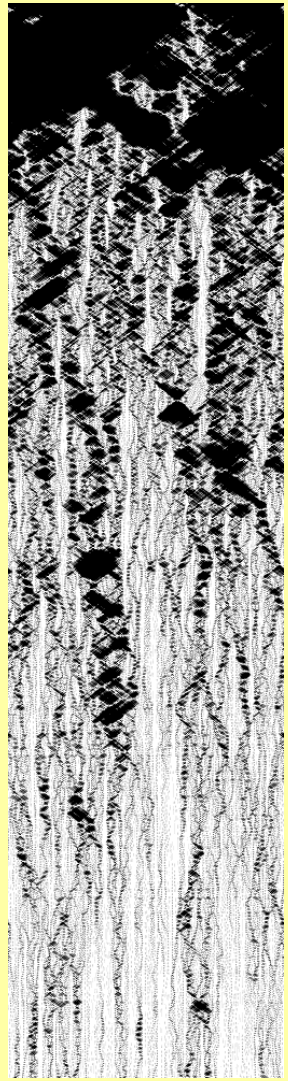
$\Rightarrow$ at $T = 0$ “zips” the chain back

Add thermal fluctuations:

$$F_{DW} = const + D\left(\frac{k_B T}{2}\sqrt{\frac{2Da^2}{K}} - 2\right)\frac{x_0}{\ell}$$

at T_c fluctuations stabilize the DW.

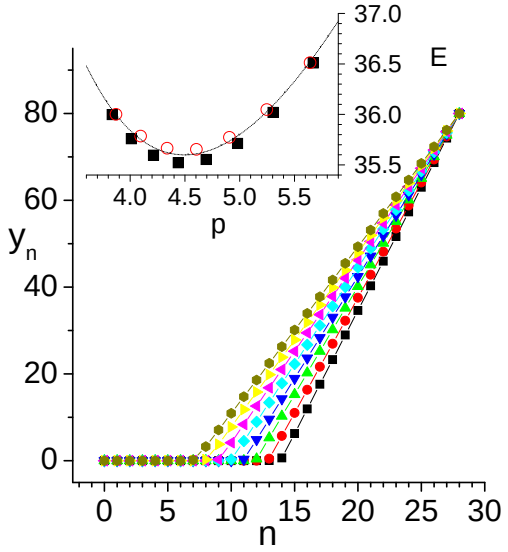
The energy of the wall is **infinite** $\Rightarrow$ allows the 1D transition.

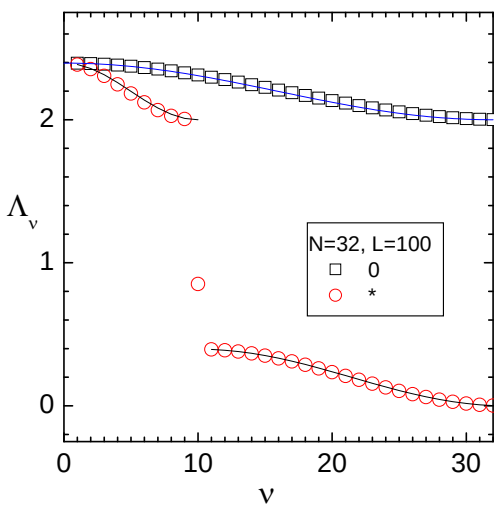


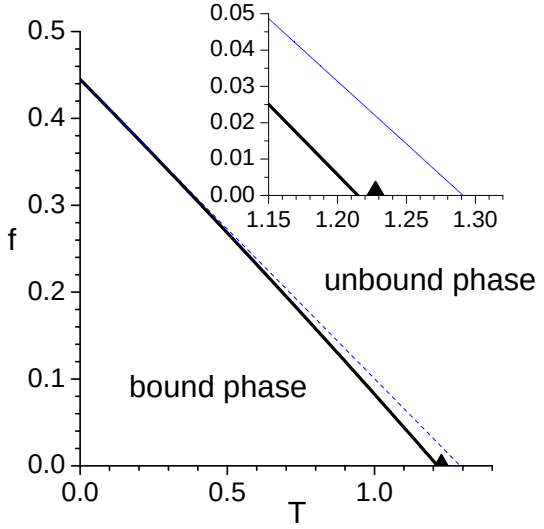
Discreteness effects on the thermodynamics of DNA melting.

N. Theodorakopoulos, M. Peyrard and R.S. Mackay, Phys. Rev. Lett. **93**, 258101 (2004)

Domain wall minimizes $\Phi = \sum_{n=0}^{N} \frac{1}{2} S(Y_{n+1} - Y_n)^2 + \left[e^{-Y_n} - 1\right]^2$







The solution is not unique (trapping by discreteness): one stable solution with minimal energy

The spectrum around the domain wall has a local mode in the gap

Free energy from the discrete spectrum → small correction to T_c

From statistics to dynamics: the limitations of the model

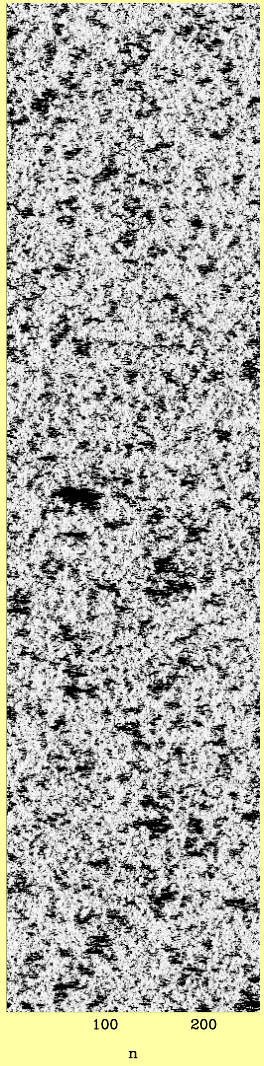
Simulation $\rightarrow$ 20 ns
Experiment:

- lifetime of an open state 30 to 300 ns
- lifetime of a closed base pair 1 to 10 ms

Nonlinear model: discrete breathers.

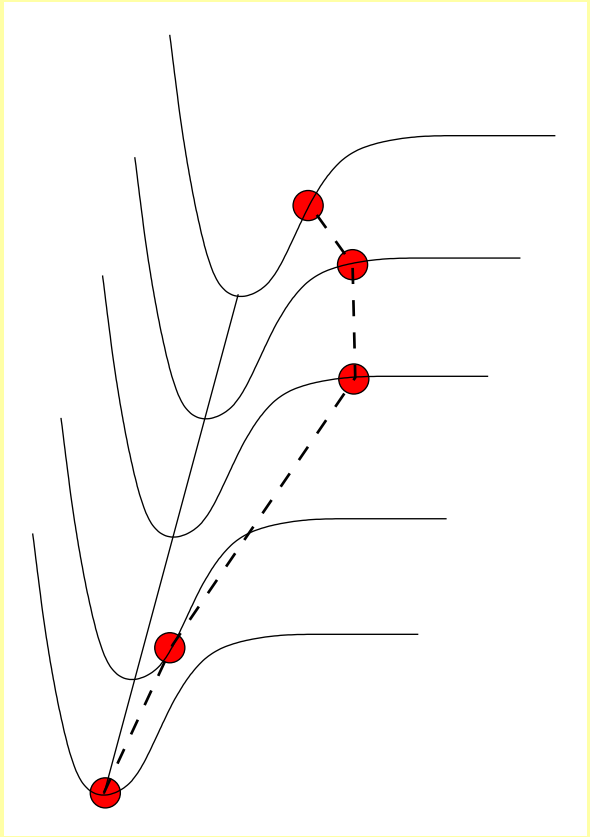
they may have a lifetime of hundreds of picoseconds (with thermal fluctuations) but opening and closing have periods of a few ps.

Disagreement by several orders of magnitude!

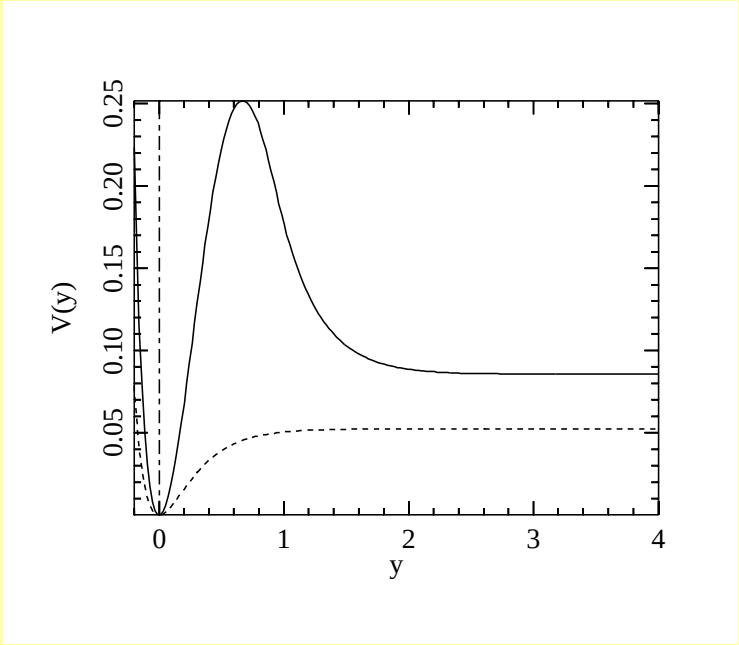


The Morse potential must be changed:

For a mesoscopic model, the “potential” is a free energy
(includes missing degrees of freedom)



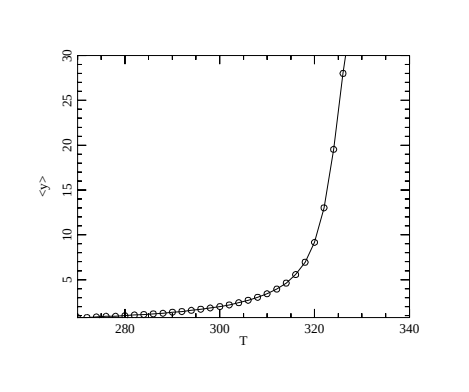
$$F = U - TS$$



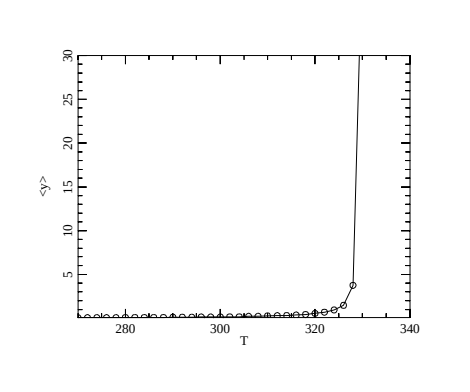
$$V_h(y) = \begin{cases} A\left[e^{-\alpha y} - 1\right]^2 & \text{if} \quad y < 0, \\ ay^2 + by^3 + cy^4 & \text{if} \quad 0 \le y \le 1, \\ D + Ee^{-\beta y}\left(y + \frac{1}{\beta}\right) & \text{if} \quad y > 1\ , \end{cases}$$

The equilibrium properties are hardly affected (the melting transition gets sharper)

Morse

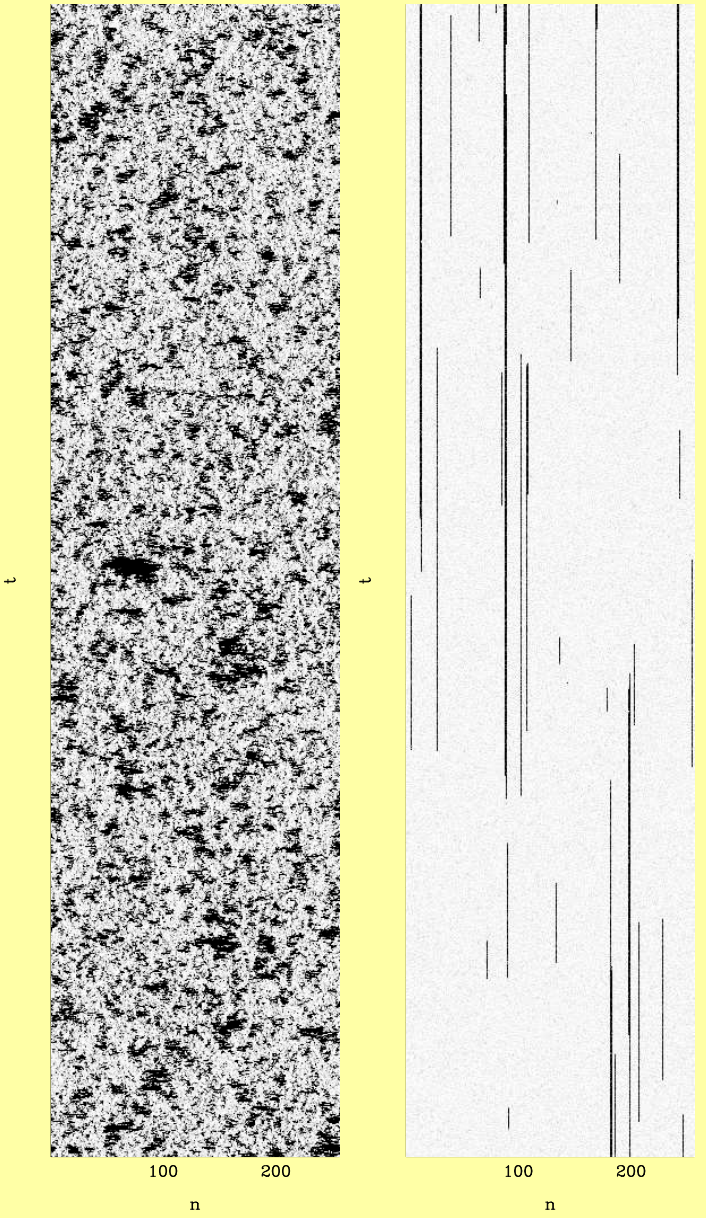


Modified Morse



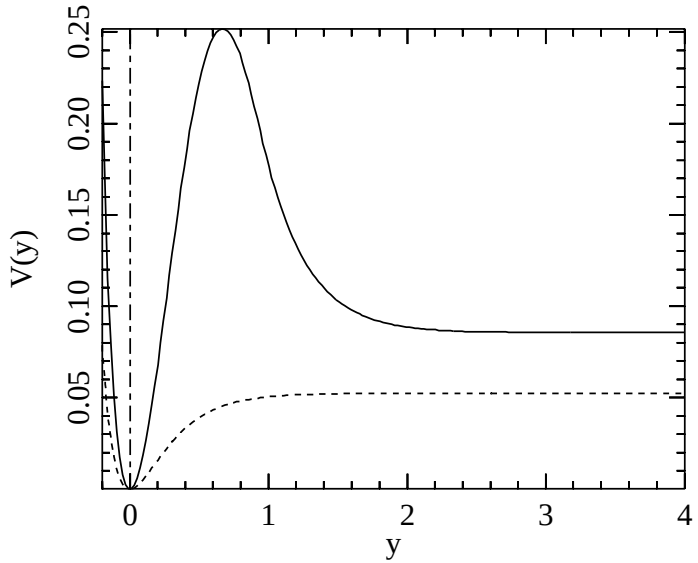
The dynamics is drastically changed:

	Morse	Modified M.
Lifetime open	0.08 ns	7 ns
Lifetime closed	0.4 ns	$0.4\ \mu$s



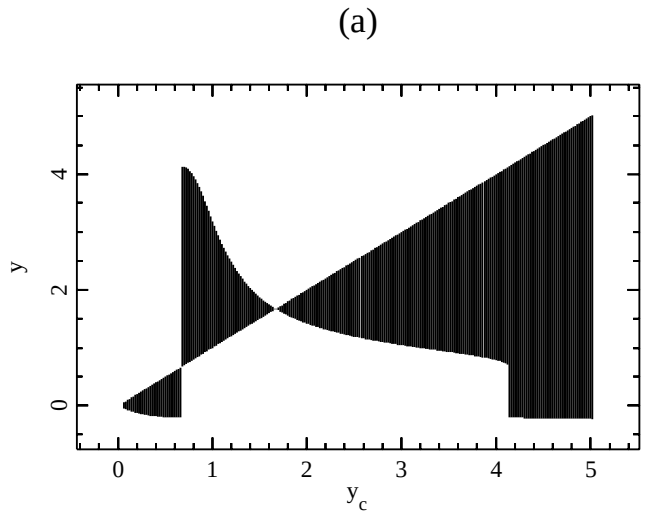
A new class of discrete breathers

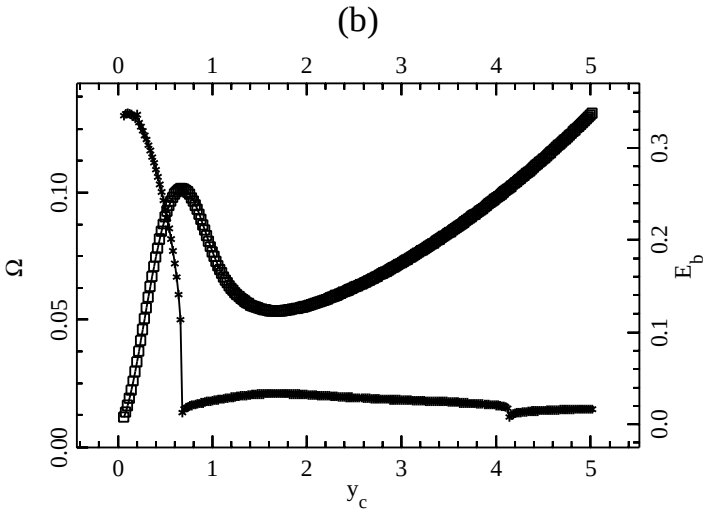
$$H = \sum_{n=-N/2}^{N/2} \frac{p_n^2}{2m} + \frac{1}{2}K(y_n - y_{n-1})^2 + V_h(y_n)$$



Anticontinuum limit cannot be used

Amplitude

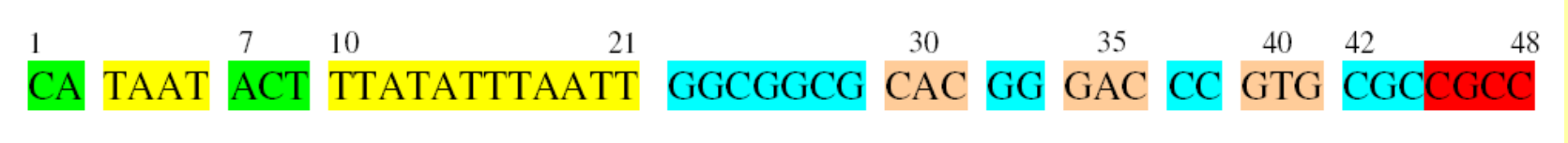




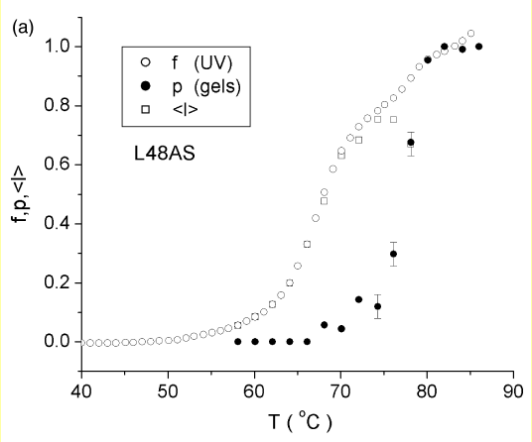
Frequency and energy

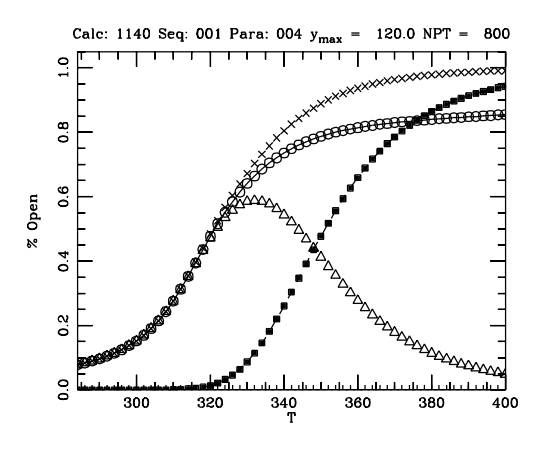
The effect of the sequence: a new challenge

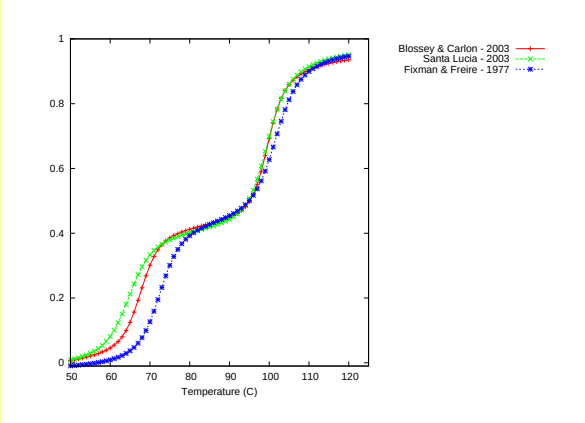
(work in progress)

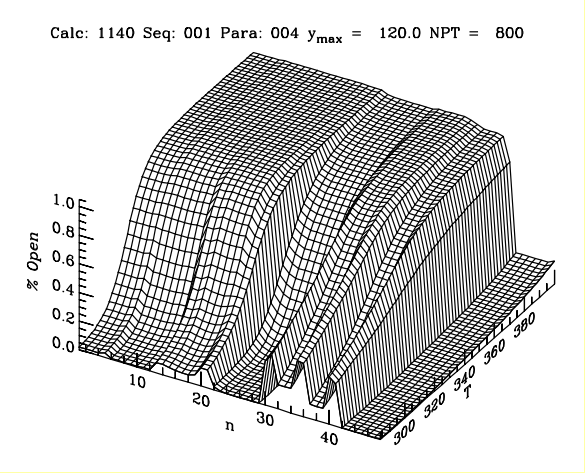


Y. Zeng, A. Montrichok, G. Zocchi, J. Mol. Biol. **339** 67-75 (2004)









Conclusion

Why does a physicist study DNA?

- Of course fascinating topic !
- It allows experiments that would be unthinkable in a “purely physical system”, thanks to the tools developed by biologists.
- It raises fundamental questions such as the phase transition in 1D, localization of thermal fluctuations, non local sequence effects . . .

It should attract the attention of mathematicians too!

- Mechanism for nonlinear localization by breather collisions (define a global measure and calculates its time evolution ?)
- New model with a breather that oscillates around a point which is unstable in the anticontinuum limit
- The phase transition in one dimension (diverging partition function, nature of the transition)