I Questions on the lectures

1. $$\mathbb{E}[Z_N] = e^{Na}\,\mathbb{E}\left[e^{\sqrt{N}X}\right] = e^{Na}\,e^{\frac{1}{2}(\sqrt{N})^2\,\mathbb{E}[X^2]}$$ because X is Gaussian with zero mean

$$\frac{1}{N}\ln \mathbb{E}[Z_N] = a + \frac{1}{2}b \qquad (i)$$

$$\mathbb{E}[\ln Z_N] = \mathbb{E}\left[Na + \sqrt{N}X\right] = Na$$

$$\frac{1}{N}\mathbb{E}[\ln Z_N] = a \qquad (ii)$$

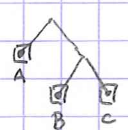
(i) is an annealed average, larger than the quenched average (ii) (as it should according to Jensen inequality). The small fluctuations of order $\sqrt{N}$ of $\ln Z_N$ around its typical value are amplified exponentially in the annealed average, that differs from the quenched one because of rare positive fluctuations.

2. The entropy of the law p is, in bits, $-p_A\log_2 p_A - p_B\log_2 p_B - p_C\log_2 p_C = \frac{1}{2}\log_2(2) + 2\,\frac{1}{4}\log_2(4)$
$= \frac{3}{2}$

Hence the compressed string is at least $1000 \times \frac{3}{2} = 1500$ bits long

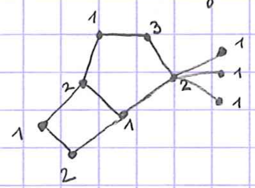
The most frequent symbol (A) should be given a shorter encoding than B and C.

For instance
$A \to 0$
$B \to 10$
$C \to 11$



has average length $\frac{1}{2}\cdot 1 + 2\cdot\frac{1}{4}\cdot 2 = \frac{3}{2}$, which is optimal

3. One cannot color this graph properly with 2 colors because of the pentagon: any cycle of odd length requires at least three colors. With 3 colors, once the pentagon is colored it is easy to complete the coloring on the other vertices of the graph, for instance



4. $$F_{\hat{M}_n}(x) = \mathbb{P}\left[\frac{\max(X_1,\dots,X_n)-a_n}{b_n} \leq x\right] = \mathbb{P}\left[\max(X_1,\dots,X_n) \leq a_n + x\,b_n\right] = \left(F_X(a_n + x\,b_n)\right)^n$$

for this to converge to a non trivial limit one needs to choose a_n and b_n such that

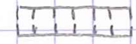
$$F_X(a_n + x\,b_n) = 1 - \frac{\gamma(x)}{n} + o\left(\frac{1}{n}\right) \qquad \text{here} \quad \exp\left(-(a_n + x\,b_n)^\beta\right) = \frac{\gamma(x)}{n} + o\left(\frac{1}{n}\right)$$

$$a_n + x\,b_n = \left(\ln n - \ln\gamma(x)\right)^{1/\beta} = (\ln n)^{1/\beta}\left(1 - \frac{\ln\gamma(x)}{\ln n}\right)^{1/\beta} \sim (\ln n)^{\frac{1}{\beta}} - \frac{1}{\beta}(\ln n)^{\frac{1}{\beta}-1}\ln\gamma(x)$$

with $a_n = (\ln n)^{\frac{1}{\beta}}$ and $b_n = \frac{1}{\beta}(\ln n)^{\frac{1}{\beta}-1}$ one has the convergence of $\hat{M}_n$ to a Gumbel distribution

II The denaturation transition of DNA

1. The homogeneous case

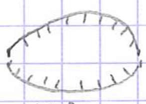
1. a) for the configuration  $k = L$, $(\ell_1, \dots, \ell_k) = (1, 2, \dots, L)$,

hence the weight is $A(1)^{L-1} \left(e^{-\beta\varepsilon}\right)^L$

The partition function being a sum of positive terms, $Z_L^h(\beta,\varepsilon) \geqslant A(1)^{L-1}\left(e^{-\beta\varepsilon}\right)^L$

$$-\frac{1}{\beta L} \ln Z_L^h(\beta,\varepsilon) \leqslant -\frac{1}{\beta L} \ln\left(A(1)^{L-1}\left(e^{-\beta\varepsilon}\right)^L\right) \xrightarrow[L\to\infty]{} \varepsilon - \frac{1}{\beta}\ln A(1)$$

$$\boxed{f^h(\beta,\varepsilon) \leqslant \varepsilon - \frac{1}{\beta}\ln A(1)}$$

b) for the configuration  $k=2$, $(\ell_1, \ell_2) = (1, L)$

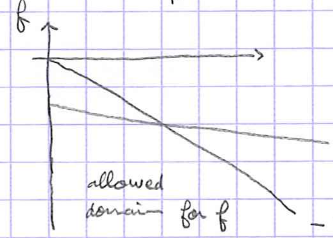
the weight is $A(L-1)\left(e^{-\beta\varepsilon}\right)^2$

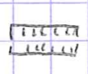
$Z_L^h(\beta,\varepsilon) \geqslant A(L-1)\left(e^{-\beta\varepsilon}\right)^2$, $\quad -\frac{1}{\beta L}\ln Z_L^h(\beta,\varepsilon) \leqslant -\frac{1}{\beta L}\ln\left(A(L-1)\left(e^{-\beta\varepsilon}\right)^2\right)$

that tends to $-\frac{1}{\beta}\ln\Omega$ when $L\to\infty$ given the assumptions on $A(\ell)$ at large ℓ

$$\boxed{f^h(\beta,\varepsilon) \leqslant -\frac{1}{\beta}\ln\Omega}$$

c) As a function of $T = \frac{1}{\beta}$, with $\varepsilon < 0$ and $\Omega > A(1)$ these two bounds cross,



when $T \to 0$ the energetic effect dominates, the relevant bound is indeed due to 

When $T \to \infty$ the entropic effect dominates, 

2. $$Z^h(\beta,\varepsilon) = \sum_{k=2}^{L}\left(e^{-\beta\varepsilon}\right)^k \sum_{1<\ell_2<\dots<\ell_{k-1}<L} A(\ell_2-\ell_1)\dots A(\ell_k-\ell_{k-1})$$

$$\frac{\partial}{\partial\varepsilon} Z^h(\beta,\varepsilon) = -\beta\sum_{k=2}^{L} k \left(e^{-\beta\varepsilon}\right)^k \dots$$

$$\frac{\partial}{\partial\varepsilon}\ln Z^h(\beta,\varepsilon) = \frac{\frac{\partial}{\partial\varepsilon} Z^h(\beta,\varepsilon)}{Z^h(\beta,\varepsilon)} = -\beta\langle k\rangle$$, $\langle k\rangle$ is the average number of contacts in the molecule

Hence Θ is the average fraction of contact in the thermodynamic limit if $\theta = 0$ (at high temperature) the molecule is essentially detached, at low enough temperatures $\theta > 0$, a finite fraction of the bases are in contact

3. a) At the leading exponential level $Z_L^h(\beta,\varepsilon) \sim e^{-\beta L f^h(\beta,\varepsilon)}$,

hence $Z_L^h(\beta,\varepsilon)\, x^L \sim \left(x\, e^{-\beta f^h(\beta,\varepsilon)}\right)^L$

the series converges if $x e^{-\beta f^h(\beta,\varepsilon)} < 1$
——— diverges ——— > 1
$\Rightarrow x_*(\beta,\varepsilon)\, e^{-\beta f^h(\beta,\varepsilon)} = 1$

$$\boxed{f^h(\beta,\varepsilon) = \frac{1}{\beta} \ln x_*(\beta,\varepsilon)}$$

b) $$Z^h(x,\beta,\varepsilon) = \sum_{L=2}^{\infty} x^L \sum_{k=2}^{L} \left(e^{-\beta\varepsilon}\right)^k \sum_{1 < i_2 < \dots < i_{k-1} < L} A(i_2 - i_1) \cdots A(i_k - i_{k-1})$$ with $i_1 = 1$, $i_k = L$

$$= \sum_{k=2}^{\infty} \left(e^{-\beta\varepsilon}\right)^k \sum_{1 < i_2 < i_3 < \dots < i_k} A(i_2-i_1) \dots A(i_k - i_{k-1})\; x^{i_1} x^{i_2-i_1} \dots x^{i_k - i_{k-1}}$$

$$= x \sum_{k=2}^{\infty} \left(e^{-\beta\varepsilon}\right)^k \left(\sum_{\ell=1}^{\infty} A(\ell)\, x^\ell\right)^{k-1}$$ changing independent variables to $i_2 - i_1$, $i_3 - i_2$, …

$$= x \left(e^{-\beta\varepsilon}\right)^2 \hat{A}(x) \sum_{k=0}^{\infty} \left(e^{-\beta\varepsilon} \hat{A}(x)\right)^k$$

$$\boxed{Z^h(x,\beta,\varepsilon) = \frac{x\left(e^{-\beta\varepsilon}\right)^2 \hat{A}(x)}{1 - e^{-\beta\varepsilon}\hat{A}(x)}}$$ if $e^{-\beta\varepsilon}\hat{A}(x) < 1$ for the geometric series to converge

one thus sees that $x_*(\beta,\varepsilon) = \sup\left\{x : e^{-\beta\varepsilon}\hat{A}(x) < 1\right\}$, hence

$$\boxed{f^h(\beta,\varepsilon) = \frac{1}{\beta} \ln\left(\sup\left\{x : e^{-\beta\varepsilon}\hat{A}(x) < 1\right\}\right)}$$

4. a) $\hat{A}(0) = 0$ as the series start with $\ell = 1$

b) $\hat{A}(x) = A(1)\, x + O(x^2)$ as $x \to 0$

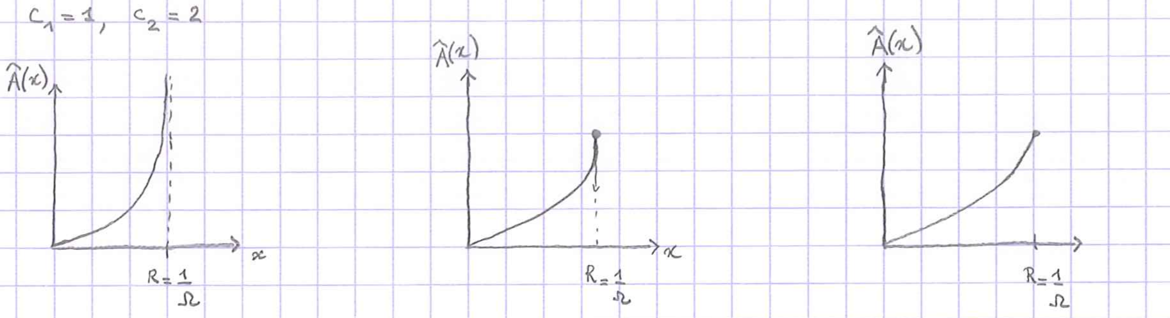
c) $\hat{A}(x) = \sum_{\ell \geq 1} A(\ell)\, x^\ell$ with $A(\ell) > 0$ $\forall \ell$, a sum of increasing functions (for $x \geq 0$) is thus increasing

d) at the leading exponential order $A(\ell) \sim \Omega^\ell$ $\Rightarrow R = \frac{1}{\Omega}$, $(x\Omega)^\ell \to 0$ for $x < R$, $\to \infty$ for $x > R$

e) in the series for $\hat{A}(R)$ one has $\sum_\ell \frac{1}{\ell^c}$, finite for $c > 1$, divergent for $c < 1$

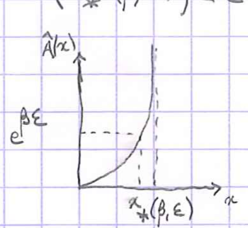
f) $\hat{A}'(x) = \sum_{\ell=1}^{\infty} A(\ell)\, \ell\, x^{\ell-1}$, in $\hat{A}'(R)$ one has $\sum_\ell \frac{1}{\ell^{c-1}}$, finite for $c > 2$, divergent for $c < 2$

$c_1 = 1, \quad c_2 = 2$



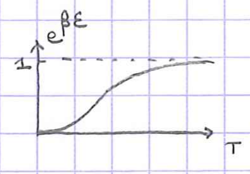
$c < 1, \quad \hat{A}(R) = \infty$ $\qquad$ $1 < c < 2, \quad \hat{A}(R) < \infty, \quad \hat{A}'(R) = \infty$ $\qquad$ $c > 2, \quad \hat{A}(R) < \infty, \quad \hat{A}'(R) < \infty$

5. as a function of the temperature $e^{-\beta\varepsilon}$ varies smoothly, $x_*(\beta,\varepsilon)$ such that $\hat{A}(x_*(\beta,\varepsilon)) = e^{\beta\varepsilon}$ as well, as there is always a solution in this case $c<1$:



And so it is also the case of $f^h(\beta,\varepsilon) = \frac{1}{\beta} \ln x_*(\beta,\varepsilon)$

6. Remembering that $\varepsilon < 0$, as a function of T the variations of $e^{\beta\varepsilon}$ are of the form:



when T is large $e^{\beta\varepsilon}$ becomes close to 1, if $\hat{A}(R) < 1$ there is no more solution to the equation $\hat{A}(x) = e^{\beta\varepsilon}$ on the interval $[0,R]$, in that case $x_*(\beta,\varepsilon) = R = \frac{1}{\Omega}$ $\Rightarrow$ $\boxed{f^h(\beta,\varepsilon) = -T \ln \Omega \text{ for } T > T_c}$, which saturates the bound of question 1. b)

As f is independent of ε, $\boxed{\theta(\beta,\varepsilon) = 0 \text{ for } T > T_c}$, ie the molecule is detached, the typical configurations have a sub-extensive number of contacts

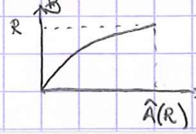
the critical temperature is such that $\hat{A}(R) = e^{\beta_c \varepsilon}$

$$\beta_c \varepsilon = \ln \hat{A}(R), \quad \beta_c = \frac{\ln \hat{A}(R)}{\varepsilon}, \quad \boxed{T_c = \frac{\varepsilon}{\ln \hat{A}(R)}}$$

7. a) when $T < T_c$ one has $e^{\beta\varepsilon} < \hat{A}(R)$, hence $x_*(\beta,\varepsilon)$ is a non-trivial solution of $\hat{A}(x) = e^{\beta\varepsilon}$, ie $x_*(\beta,\varepsilon) = B(e^{\beta\varepsilon})$ with B the reciprocal of $\hat{A}$

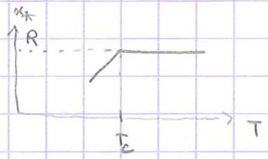
This gives $\boxed{f^h(\beta,\varepsilon) = \frac{1}{\beta} \ln \left(B(e^{\beta\varepsilon})\right)}$

b) The graph of B is obtained from the one of $\hat{A}$ by exchanging the horizontal and vertical axes, hence



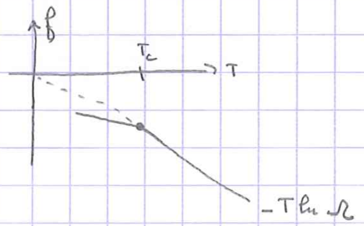
when $c > 2$, with a non zero slope at the end-point of the curve

As a function of T, around T_c x_* has the shape



ie it is continuous with a discontinuity in the derivative

Hence $f = \frac{1}{T} \ln x_*$:



is also continuous with a discontinuity in its derivative $\Rightarrow$ first order transition at T_c

The order parameter Θ is zero for $T > T_c$, and jumps to a strictly positive value (proportional to the derivative of B in $\hat{A}(R)$) for $T \to T_c^-$

c) for $1<c<2$, $\hat{A}(x)$ has a singularity when $x \to R^-$, because $\hat{A}'(x)$ diverges. To estimate this divergence we can keep the large ℓ, small $R-x$ terms only, and forget multiplicative constants:

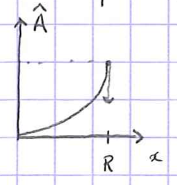
$$\hat{A}'(x) \simeq \sum_\ell \left(\frac{x}{R}\right)^\ell \frac{1}{\ell^{c-1}} \simeq \sum_\ell \left(1 - \frac{R-x}{R}\right)^\ell \frac{1}{\ell^{c-1}} \simeq \sum_\ell e^{-\ell \frac{R-x}{R}} \frac{1}{\ell^{c-1}}$$

$$\simeq \int_{\ell_0}^\infty d\ell \frac{1}{\ell^{c-1}} e^{-\ell \frac{R-x}{R}} \underset{u = \ell \frac{R-x}{R}}{\simeq} \left(\frac{R-x}{R}\right)^{c-2} \underbrace{\int_{\ell_0 \frac{R-x}{R}}^\infty du \, e^{-u} u^{1-c}}_{\text{converges when } x \to R}$$

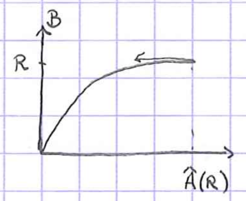
hence $\hat{A}'(x) = \text{cst} \, (R-x)^{c-2} + o\left((R-x)^{c-2}\right)$ when $x \to R^-$

integrating, $\hat{A}(x) \sim \hat{A}(R) - \text{cst} \, (R-x)^{c-1}$ when $x \to R^-$

for the reciprocal,

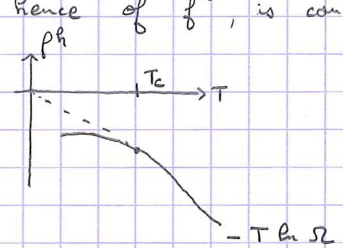


$\Leftrightarrow$



$$B\left(\hat{A}(R) - \delta\right) \sim R - \text{cst} \, \delta^{\frac{1}{c-1}}$$

Now the derivative of x_*, and hence of f^h, is continuous at T_c, hence the transition is second order.



the singular behavior of f^h for $T \to T_c^-$ is thus $\text{cst} \, (T_c - T)^{\frac{1}{c-1}}$,

that one writes with the usual α exponent as $(T_c - T)^{2-\alpha}$,

$$2 - \alpha = \frac{1}{c-1} \Leftrightarrow \alpha = \frac{2c-3}{c-1}$$

8. What are forbidden are phase transitions for 1d systems with finite range interactions. In this model the entropic term $A(\ell)$ provides an effective long range interaction, for which phase transitions in 1d are not impossible

2. The disordered case

1. $$Z_L(\beta, \underline{\varepsilon}) = \sum_{k=2}^{L} \sum_{1<\ell_2<\dots<\ell_{k-1}<L} A(\ell_2-\ell_1)\dots A(\ell_k-\ell_{k-1}) \; e^{-\beta\varepsilon_{\ell_1}} e^{-\beta\varepsilon_{\ell_2}} \dots e^{-\beta\varepsilon_{\ell_k}}$$

$$\mathbb{E}\left[Z_L(\beta,\underline{\varepsilon})\right] = \sum_{k=2}^{L} \quad '' \quad '' \quad '' \quad \left(\mathbb{E}\left[e^{-\beta\varepsilon}\right]\right)^k$$ as the energies are iid

$$= Z_L^h(\beta, \varepsilon_{eff})$$ if one defines the effective energy by

$$\boxed{e^{-\beta\varepsilon_{eff}} = \mathbb{E}\left[e^{-\beta\varepsilon}\right], \text{ or equivalently} \quad \varepsilon_{eff} = -\frac{1}{\beta}\ln \mathbb{E}\left[e^{-\beta\varepsilon}\right]}$$

2. From Jensen's inequality the quenched free energy is lower bounded by the annealed one, $f_q(\beta) \geq -\frac{1}{\beta}\lim_{L\to\infty}\frac{1}{L}\ln \mathbb{E}\left[Z_L(\beta,\underline{\varepsilon})\right] = f^h(\beta, \varepsilon_{eff})$

$$\boxed{f_q(\beta) \geq f^h(\beta, \varepsilon_{eff})}$$

$T_{c,a}$, the annealed critical temperature, corresponds to the critical temperature of the effective homogeneous model

$$f^h(\beta, \varepsilon_{eff}) = \begin{cases} -T\ln\Omega & \text{when } e^{\beta\varepsilon_{eff}} > \hat{A}(R) \\ \frac{1}{\beta}\ln\left(B(e^{\beta\varepsilon_{eff}})\right) & \text{when } e^{\beta\varepsilon_{eff}} < \hat{A}(R) \end{cases}$$

$$e^{\beta\varepsilon_{eff}} = \frac{1}{\mathbb{E}\left[e^{-\beta\varepsilon}\right]}$$ is increasing with T, the critical temperature of the annealed bound is such that $\boxed{\mathbb{E}\left[e^{-\frac{1}{T_{c,a}}\varepsilon}\right] = \frac{1}{\hat{A}(R)}}$

3. Even in presence of disorder one can write $Z_L(\beta,\underline{\varepsilon}) \geq A(L-1)\, e^{-\beta(\varepsilon_1+\varepsilon_L)}$

$$-\frac{1}{\beta L}\ln Z_L(\beta,\underline{\varepsilon}) \leq -\frac{1}{\beta L}\ln A(L-1) + \frac{\varepsilon_1+\varepsilon_L}{L}$$

Taking average and $L\to\infty$ shows that $\boxed{f_q(\beta) \leq -T\ln\Omega}$

4. For $T \geq T_{c,a}$ we have shown $\left.\begin{matrix} f_q(\beta) \geq -T\ln\Omega \\ \text{and } f_q(\beta) \leq -T\ln\Omega \end{matrix}\right\} \Rightarrow f_q(\beta) = -T\ln\Omega$

In the high temperature phase the molecule is detached, there are no binding energy effects, hence whether $\underline{\varepsilon}$ is disordered or homogeneous should not have any effect

5. For $T > T_{c,a}$ $f_q(\beta) = -T\ln\Omega$ is perfectly analytic : if f_q has a singularity for some temperature $T_{c,q}$ it is certainly at a lower temperature

3. <u>Harris criterion</u>

1. $$Z_L(\beta,\underline{\varepsilon}) = \sum_{k=2}^{L} \sum_{1<\ell_2<\dots<\ell_{k-1}<L} A(\ell_2-\ell_1)\dots A(\ell_k-\ell_{k-1}) \left(e^{-\beta\varepsilon}\right)^k \quad (1+u_{\ell_1})(1+u_{\ell_2})\dots(1+u_{\ell_k})$$

2. $(1+u_{\ell_1})\dots(1+u_{\ell_k}) = 1 + (u_{\ell_1}+\dots+u_{\ell_k}) + O(u^2)$, hence

$$Z_L(\beta,\underline{\varepsilon}) = Z_L^h(\beta,\varepsilon) + \sum_{\ell=1}^{L} u_\ell \underbrace{\sum_{k=2}^{L} \sum_{1<\ell_2<\dots<\ell_{k-1}<L} A(\ell_2-\ell_1)\dots A(\ell_k-\ell_{k-1}) \left(e^{-\beta\varepsilon}\right)^k \left(\delta_{\ell_1,\ell}+\delta_{\ell_2,\ell}+\dots+\delta_{\ell_k,\ell}\right)}_{Z_L^h(\beta,\varepsilon)\,\times\, p_\ell(L,\beta,\varepsilon)},$$

where $p_\ell(L,\beta,\varepsilon)$ is the probability, in the homogeneous model, of the event "the base at position ℓ is in contact"

3. $\ln(1+x) = x - \frac{x^2}{2} + O(x^3)$

$$\ln Z_L(\beta,\underline{\varepsilon}) = \ln Z_L^h(\beta,\varepsilon) + \ln\left(1 + \sum_\ell u_\ell p_\ell\right)$$

$$\text{"} \quad = \quad \text{"} \quad + \sum_\ell u_\ell p_\ell - \frac{1}{2}\sum_{\ell,\ell'} u_\ell u_{\ell'} p_\ell p_{\ell'}$$

$$\mathbb{E}\left[\ln Z_L(\beta,\underline{\varepsilon})\right] = \ln Z_L^h(\beta,\varepsilon) - \frac{1}{2}\mathbb{E}[u^2]\sum_{\ell=1}^{L} p_\ell^2$$

as $\mathbb{E}[u_\ell] = 0$

$\mathbb{E}[u_\ell u_{\ell'}] = \mathbb{E}[u^2]\,\delta_{\ell,\ell'}$ by their independence on different sites

$$-\frac{1}{\beta L}\mathbb{E}\left[\ln Z_L(\beta,\underline{\varepsilon})\right] = -\frac{1}{\beta L}\ln Z_L^h(\beta,\varepsilon) + \frac{1}{2\beta}\mathbb{E}[u^2]\left(\frac{1}{L}\sum_{\ell=1}^{L} p_\ell(L,\beta,\varepsilon)^2\right)$$

4. Apart for boundary effects when $\ell \simeq 1$ and $\ell \simeq L$, $p_\ell(L,\beta,\varepsilon) \to \theta(\beta,\varepsilon)$, the fraction of bases in contact, hence in the thermodynamic limit

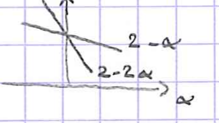
$$f_q(\beta) = f^h(\beta,\varepsilon) + \frac{1}{2\beta}\mathbb{E}[u^2]\,\theta(\beta,\varepsilon)^2$$

At $T \to T_c^-$, $\quad f^h \sim (T_c-T)^{2-\alpha}$

$\theta \sim (T_c-T)^{1-\alpha}$ (it is a derivative of f)

$\theta^2 \sim (T_c-T)^{2-2\alpha}$

The exponent of the two terms are $2-\alpha$ and $2-2\alpha$; as $(T_c-T)\to 0$, the dominant term is the one with the minimal exponent.



For $\alpha>0$, for any arbitrary small disorder, sufficiently close to T_c the disorder term dominates, hence the limit disorder $\to 0$ cannot be smooth, the disorder is relevant and modifies the critical behavior of the homogeneous model