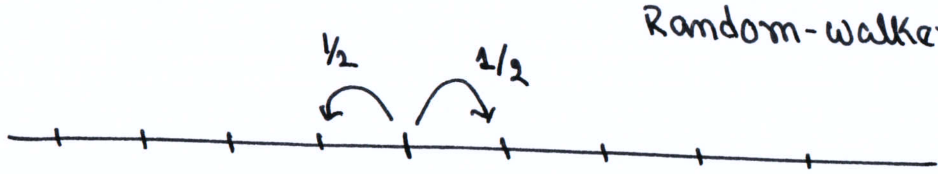
Beyond ~~and~~ the realm of central-limit-theorem.

Q. How good is CLT ?

Random-walker



Position after n-steps

$$M_N = \sum_{i=1}^{n} x_i \qquad x_i = \left\{ \pm 1 \text{ with prob } \tfrac{1}{2} . \right.$$

Exact expression for probability

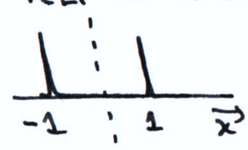
$$P_{Ex}(M_N) = \frac{1}{2} \cdot \frac{1}{2^n} \binom{N}{\frac{N+M}{2}} \qquad \text{[Binomial distribution]}$$

why is this ?

CLT gives

$$P_{CLT}(M_N) \simeq \frac{1}{\sqrt{2\pi N}} e^{-\frac{M_N^2}{2N}}$$

Remark: check how large N is needed for P_{CLT} to match well.
Is it 10^3 or 10 ? Note. $p(x)$



How well does CLT compare with exact P ?

For tail-events, say $M_N = N$

$$P_{CLT}(N) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{N}{2}} , \qquad P_{Ex}(N) = \frac{1}{2^{N+1}} \simeq e^{-N \ln 2}$$

Relative error: $\dfrac{P_{CLT} - P_{EX}}{P_{EX}} \simeq e^{N(\ln 2 - \frac{1}{2})} \simeq e^{0.2 N}$

Error is exponentially large !

** CLT describes typical fluctuations ($\sim\sqrt{N}$), but fails for rare events **

Large deviation theory describes better. [Very important when we discuss non-eq stat mech, including quantum.]

$$P(M) = \frac{1}{2^{N+1}} \cdot \frac{N!}{\left(\frac{N+M}{2}\right)!\left(\frac{N-M}{2}\right)!}$$

use stirlings approx $N \gg 1$ and $M \gg 1$

$$\simeq e^{-\frac{N}{2}\left\{\left(1+\frac{M}{N}\right)\ln\left(1+\frac{M}{N}\right)+\left(1-\frac{M}{N}\right)\ln\left(1-\frac{M}{N}\right)\right\}}$$

for $M \gg 1$ and $N \gg 1$.

Case 1: $M = m \cdot N^{\alpha}$ with $\alpha < 1$ $\Rightarrow$ $\frac{M}{N} \ll 1$ for large N.

$$P(M) \simeq A \cdot e^{-\frac{N}{2}\cdot\left(\frac{M}{N}\right)^2}$$

$$\simeq A \cdot e^{-\frac{M^2}{2N}} \qquad \Leftarrow \text{CLT is good!}$$

Case 2: $M = m \cdot N$

$$\boxed{P(M = mN) \simeq A \cdot e^{-N\phi(m)}}$$

with $\phi(m) = \frac{1}{2}\left\{(1+m)\ln(1+m) + (1-m)\ln(1-m)\right\}$

Case 3:

$$P(M = m N^{\alpha}) \sim \delta(m) \quad \text{for } N \text{ large.}$$

with $\alpha > 1$

[see using $P_{exact}(M)$ and that $(-n)! = \infty$]

Together, this implies

$$\boxed{P(M = m \cdot N^{\alpha}) \asymp e^{-N^{2\alpha-1}\,\phi_{\alpha}(m)}}$$

with

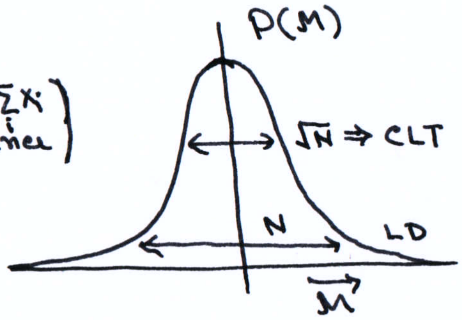
$$\phi_{\alpha}(m) = \begin{cases} \frac{1}{2} m^2 & \text{for } \alpha < 1 \\ \frac{1}{2}(1+m)\ln(1+m) + \frac{1}{2}(1-m)\ln(1-m) & \text{for } \alpha = 1 \\ \begin{cases} 0 & \text{for } m=0 \\ \infty & \text{for else} \end{cases} & \text{for } \alpha > 1. \end{cases}$$

N

→ Called a large deviation description.

$\phi_{\alpha}(m)$ is the large deviation function.

also called rate function. (think $\sum x_i$ convergence)



Remark: $\alpha = 1$ is the non-trivial case.

and often considered for LD

How well does it describe?

$$P_{LDF}(M=N) \asymp e^{-N \ln 2}$$

[take $m \to 1$ limit and $\lim\limits_{x\to 0} x \ln x = 0$]

$$P_{Ex}(M=N) \asymp e^{-N \ln 2}$$

$\boxed{\text{large deviation description is an improved version of CLT.}}$

Remark: The $A \asymp B$ sign means

$$\lim_{N\to\infty} \frac{\ln A}{\ln B} = 1.$$

- essentially means ignoring subleading terms.
- Pre-factor can be determined by normalization.

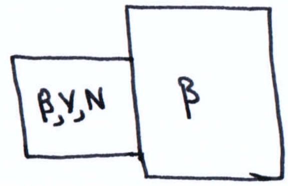
Large deviation function in equilibrium

[why do we care about it?]

(things that you all know!)

micro-canonical

Canonical



Thermodynamic potential:

entropy: $S(E,V,N) = \ln \Omega(E,V,N)$

Free-Energy: $F(\beta,V,N) = -\ln Z(\beta,V,N)$

Ensemble equivalence $(V \to \infty)$

$$F(\beta) = \max_E \{\beta E - S(E)\}$$ Legendre transform.

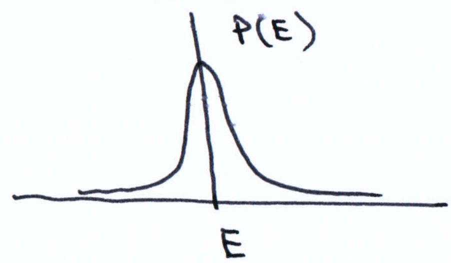
$$\Rightarrow \quad f(\beta) = \max_e \{\beta e - s(e)\}$$

$\boxed{f = \frac{1}{V}F \; ; \; s = \frac{1}{V}S}$

F and S are extensive quantities.

An Example: Let's take Canonical Ensemble

β held fixed. Energy E fluctuates



Average: $\langle E \rangle = \frac{d}{d\beta} F(\beta) = V \cdot e$

variance: $\langle E^2 \rangle_c = -\frac{d^2}{d\beta^2} F''(\beta) = V \cdot (T^2 c_v)$ ← central limit theorem! static fluctuation dissipation ~~theorem~~ relation.

Full distribution: Does stat mech give you?

$$P_{V,N}(E) = \frac{\Omega(E)\, e^{-\beta E}}{Z(\beta)} = e^{S(E) - \beta E + F(\beta)}$$

$$\asymp e^{-V \underbrace{\{\beta e - s(e) - f(\beta)\}}_{\phi(e)}}$$

$$\Rightarrow \boxed{P\left(\frac{E}{V} = e\right) \asymp e^{-V\phi(e)}}$$ large-deviation.

Simplify the formula for $\phi(e)$:

- $$f(\beta) = \max_{e} \{\beta e - s(e)\}$$

$$= \beta e^* - s(e^*) \quad \text{with} \quad \beta = s'(e^*) \iff e^* = f'(\beta)$$

Property of Legendre transform.

Gives $$\boxed{\phi(e) = s(e^*) - s(e) + \beta(e - e^*)}$$

What is e^*? $\phi(e^*) = 0 \Rightarrow e^*$ is most probable / average value of e.

See it agrees with thermodynamics. $\langle e \rangle = e^* = \frac{d}{d\beta} f(\beta)$.

$$\cancel{s(e) = \max_{\beta} \{\beta e - f(\beta)\}}$$

$$\cancel{= \beta e - f(\beta) \text{ with}}$$

⊛ Large deviation is NOT the thermodynamic free energy, but it is related to it.

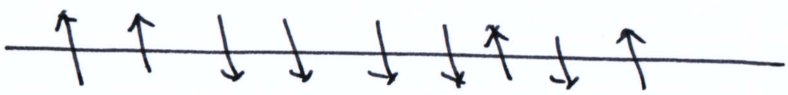
For $\alpha = 1$ case

$$\boxed{P(M = mN) \asymp e^{-N\phi(m)}}$$

You have and will encounter this form many times in physics!

Example 0: see page 19a, 19b.

Example 1.



$$P\left(\frac{M}{N} = m\right) \asymp e^{-F(m)} \leftarrow \text{Landue free energy.}$$

[What's the difference from thermodynamic free energy]

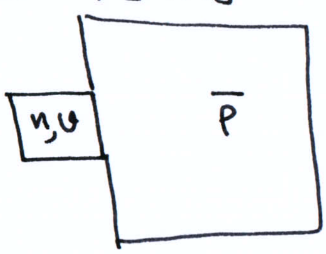
$$F(m) = N \cdot \phi(m)$$

$\uparrow$ extensivity.

LDF form works even when spins are interacting!

(in general $F[m(x)] \asymp N\,\phi[m(x)]$

Example 2: Grand canonical



$$P\left(\frac{n}{v} = \rho\right) \asymp e^{-v\phi(\rho)}$$

with

$$\phi(\rho) = \beta \left\{ f(\rho) - f(\bar{\rho}) - f'(\bar{\rho})(\rho - \bar{\rho}) \right\}$$

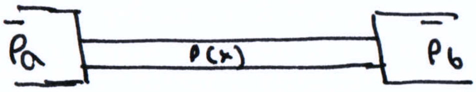
when correlations are short-ranged!

$$f(\rho) = -\frac{1}{\beta V} \ln Z \equiv \text{thermodynamic free energy.}$$

Assignment: Generalize for a profile $\rho(x)$.

$$\phi[\rho(x)] = \beta \int dx \left\{ f(\rho(x)) - f(\bar{\rho}) - f'(\bar{\rho})(\rho(x) - \bar{\rho}) \right\}$$

Example 3 :



Non-equilibrium

$$P[\rho(x)] \asymp e^{-L\cdot\phi(\rho(x))}$$

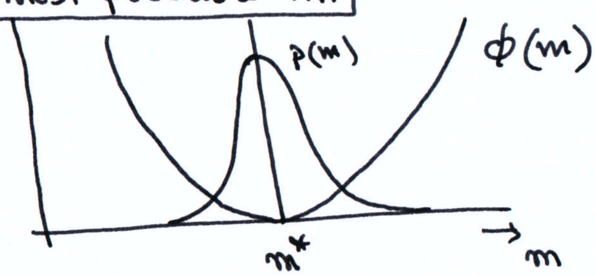
- Gives a mathematical way to generalize the idea of Landau free energy, out-side equilibrium.
- They will be important when we discuss non-equilibrium.

[Example. in Jarzynski relation]

~~Properties of a large deviation.~~

Properties of a large deviation function.

① $\boxed{\phi(m) \text{ is minimum at most probable } m.}$



$\Rightarrow \phi'(m^*) = 0$

and $\phi''(m^*) \geq 0.$

[Caution: there could be multiple minima. see later.]

② $\boxed{\phi(m) \text{ vanish at } m^*: \ \phi(m^*) = 0}$ ← assignment.

A naive argument:

$$P(m) = \frac{1}{Z_n} e^{-N\left[\phi(m^*) + \frac{1}{2}\phi''(m^*)(m-m^*)^2 + \dots\right]}$$

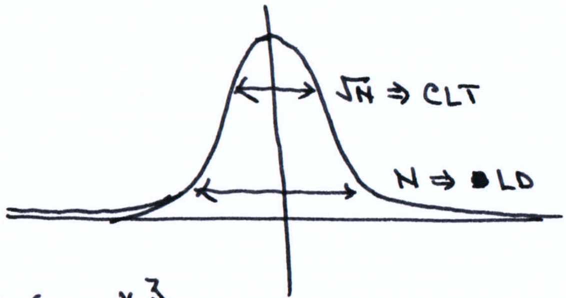
$$Z_n = e^{-N\phi(m^*)} \int dm \, e^{-\frac{N}{2}\phi''(m^*)(m-m^*)} \dots$$

cancells.

[A clean argument using cumulant generating function.]

③ Typically, CLT is recovered from ldp. (large deviation principle)

Because, typically, $\phi(m)$ is quadratic around m^*.



$$P(M = Nm) \simeq e^{-\frac{N}{2}\phi''(m^*)(m-m^*)^2 - N\cdot(m-m^*)^3\cdot * - \dots}$$

$$= e^{-\frac{(M-M^*)^2}{2N\sigma^2} - \frac{(M-M^*)^3}{N^2} * - \dots} \qquad m = \frac{M}{N}$$

higher order for large N.

$$\sigma^2 = \frac{1}{\phi''(m^*)}$$

⇒ For large N, and "small" fluctuations

$$P(M) \simeq \frac{1}{\sqrt{2\pi N\sigma^2}}\, e^{-\frac{(M-M^*)^2}{2N\sigma^2}}$$

How to calculate large deviation function

Simplest example: iid ~~and~~ (extension of CLT)

$$M_N = \sum_i x_i \qquad \text{with } \langle x \rangle \text{ and } \langle x^2 \rangle_c \text{ finite.}$$

$\lambda \in \mathrm{Re}$

$$\Rightarrow \quad G_N(\lambda) = \langle e^{\lambda M_N} \rangle \underset{\text{iid}}{=\!=} \langle e^{\lambda x_i} \rangle^N = g(\lambda)^N$$

$$\Rightarrow \quad \boxed{G_N(\lambda) = e^{N \mu(\lambda)} \qquad \ln g(\lambda)}$$

By definition

$$G_N(\lambda) = \int dM \, e^{\lambda M} P_N(M)$$

if ldf exist
$P_N(M) \approx \frac{\ast}{\sqrt{N}} e^{-N\phi(m)}$
and
$M = mN$

$$\Rightarrow \quad e^{N\mu(\lambda)} \asymp \int dm \, e^{N[\lambda m - \phi(m)]}$$

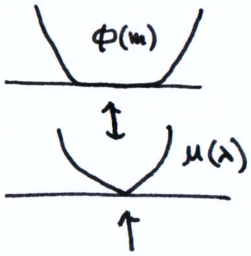
saddle point for Large N

$$\boxed{\mu(\lambda) = \max_m \{\lambda m - \phi(m)\}}$$

Legendre transform
Fenchel

if $\phi(m)$ is strictly convex ~~and~~

$$\boxed{\phi(m) = \max_\lambda \{\lambda m - \mu(\lambda)\}}$$



Singularity does not allow forward transform.

Self-duality of Legendre transform

Gives

$$\boxed{\phi(m) = \max_\lambda \{\lambda m - \log g(\lambda)\}}$$

This is how one calculates $\phi(m)$ for iid.

Cramér's Theorem.

Cramér's Theorem: for x an iid random variable, if $g(\lambda) = \langle e^{\lambda x} \rangle$ exist for Real λ, then

$$P\left(\frac{1}{n}\sum_i x_i = m\right) \simeq e^{-n\phi(m)} \quad \text{for large } n.$$

with

$$\phi(m) = \max_{\lambda} \left\{ \lambda m - \ln g(\lambda) \right\}$$

Remark: strict-convexity of $\phi(m)$ is taken care by the fact that if $g(\lambda)$ exist then it is real analytic. These kind of issues are taken care by the mathematical proof of Cramér.

Term Paper: "Making sense of the Legendre Transform" Zia et al Am J phy 77, 614 (2009)

Illustrative examples

Ex 1: $P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x^*)^2}{2\sigma^2}} \longrightarrow g(\lambda) = e^{x^*\lambda + \frac{1}{2}\sigma^2\lambda^2}$

we know (done earlier)

use Cramer's theorem

$$\phi(m) = \frac{(m - x^*)^2}{2\sigma^2}$$

both side agree.

$$P(M) = \frac{1}{\sqrt{2\pi n\sigma^2}} e^{-\frac{(M - nx^*)^2}{2n\sigma^2}} \longrightarrow P(m) \simeq \sqrt{\frac{n}{2\pi\sigma^2}}\, e^{-n\phi(m)}$$

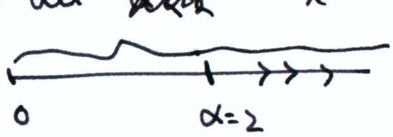
Ex 2: when does Cramér's theorem NOT work?

Ans: when $g(\lambda)$ do not exist.

Cauchy distribution.

$$P(x) = \frac{1}{\pi} \frac{c}{x^2 + c^2}$$

For all ~~αx~~ Power law $\frac{1}{x^{1+\alpha}}$



NO LDF

~~We have seen before that it is a stable distri~~

This is a power-law distribution, and

$$g(\lambda) = \langle e^{\lambda x} \rangle$$ do not exist (diverge for all real λ except $\lambda = 0$)

[note, in our earlier analysis, we showed $g(\lambda)$ exist along the imaginary line $\lambda = iK$ but non-analytic at $K = 0$.

$$g(iK) = e^{-c|K|}$$, so Wick's rotation not possible.]

means Cramér's th does not apply.

We see this as we already know that Cauchy is a stable distribution.

$$P(M) = \frac{1}{\pi} \frac{(cn)}{M^2 + (cn)^2} \sim \frac{1}{n} \quad \text{for } M = mn \ \& \ n \to \text{large}.$$

$$\not\longrightarrow \quad P(m) \sim e^{-n\phi(m)}$$

** $\boxed{\text{Cauchy does not have a large-deviation-description}}$ **

$\boxed{\text{neither CLT}}$

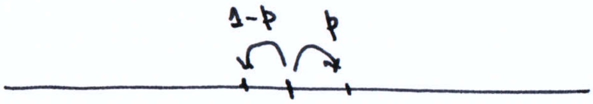
Remark: none of the power law distr. $p(x) \sim \frac{1}{x^{1+\alpha}}$ has large-deviation description ($g(\lambda)$ diverge),

$\boxed{\text{Although CLT works for } \alpha \geq 2.}$ [for an example see Exercise on next page]

Exercise: for $\alpha \geq 2$ where CLT works, estimate how good is CLT, Meaning show that CLT is good for

$$|M - n\langle x\rangle| \ll \sqrt{n \ln (n^{\alpha-2}) \cdot \langle x^2 \rangle_c}$$

Exercise:



Find the large-deviation

$$P\left(\frac{M_n}{n} = m\right) \asymp e^{-n\phi(m)}$$

① using Binomial distribution.

② Verify using Cramér's theorem.

Exercise: An example for $\alpha > 2$ where CLT works, but there is no large deviation.

$$p(x) = \frac{2c^3}{\pi (x^2 + c^2)^2} \Leftrightarrow g(ik) = (1 + c|k|)\, e^{-c|k|}$$

$$\Updownarrow$$

$$\underset{\text{Real.}}{g(\lambda)} = \begin{cases} 1 & \text{for } \lambda = 0 \\ \infty & \text{otherwise} \end{cases}$$

It can be shown that,

$$P\left(M_n = \sum_i^n x_i\right) \sim \frac{n}{M_n^4}$$

[Ref. Uchaikin et al
Chance and stability: stable distributions and their applications (1999)]

Important: Why LDF not work here?

→ [intuition: unlike in exponentially decaying distribution where each tiny x_i work together to give a large M, in power law case each single x_i can single handedly produce large M. Prob of these events is sum of each $\cdot P(x_i = M)$.]

Large-deviations can exist beyond iid's, and for general scenarios.

Varadhan

A general statement: let M_n is a random variable indexed by a real parameter. Then, if the probability

$$\mathbb{P}\left(\frac{M_n}{n^\alpha}=m\right) \asymp e^{-n^\beta \phi(m)}$$

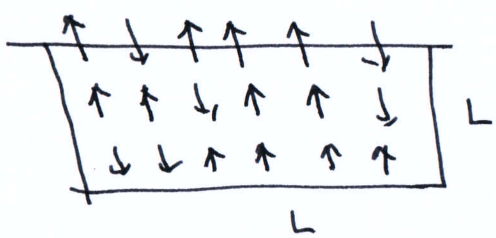
for n large with $\alpha, \beta > 0$.

exist, then it said that M_n has a large deviation description (or follows a large deviation principle) with $\phi(m)$ being the large deviation function.

A very good reference on ldf. H. Touchette, Phys Reports, 478 (2009), 1.

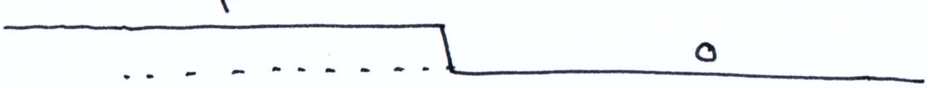
Example: Equilibrium Stat-mech

$$\mathbb{P}\left(\frac{M_L}{L^d}=m\right) \asymp e^{-L^d \phi(m)}$$



spins are correlated.

Example:



[Derrida et al, JSP (2009), 136, 1]

Q_T = net ptl flow at origin over time T.

$$\mathbb{P}\left(\frac{Q_T}{\sqrt{T}}=j\right) \asymp e^{-\sqrt{T}\cdot\phi(j)}$$

OR

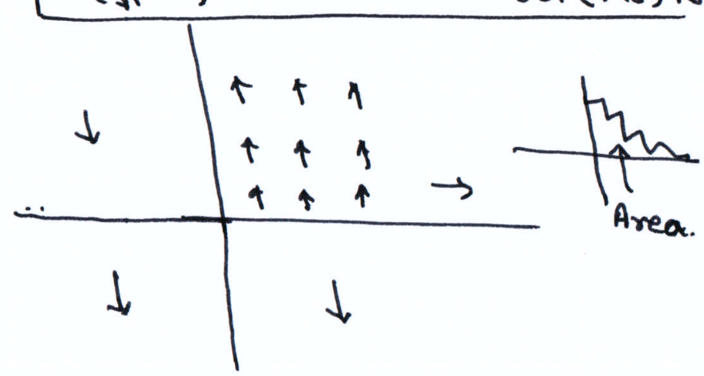
Many Brownian ptl

$$\mathbb{P}\left(\frac{X_T}{\sqrt{T}}=x\right) \asymp e^{-\sqrt{T}\,\phi(x)}$$

Krapivsky et al JSP (2015) 160, 885

Example:

$$\mathbb{P}\left(\frac{A_T}{T}=a\right) \asymp e^{-\sqrt{T}\cdot\phi(a)}$$



Ref: Melting of an Ising quadrant

Krapivsky et al. JPA 48, 015005 (2014)

How does one find large-deviation function?

For a general case (but not including all!)

<u>Gärtner-Ellis Theorem</u>: For any random variable M_n, if the <u>scaled</u> cumulant generating function

$$\mu(\lambda) = \lim_{n\to\infty} \frac{\langle e^{n\lambda m_n}\rangle}{n} \qquad \text{where } \left\{ m_n = \frac{M_n}{n} \right\}$$

<u>exist</u> and <u>differentiable</u>, then

$$P\left(\frac{M_n}{n} = m\right) \asymp e^{-n\phi(m)}$$

with

$$\phi(m) = \max_{\lambda} \{\lambda m - \mu(\lambda)\}$$

[Ref. See H. Touchette review]

<u>Remark</u>:

In comparison to Cramér's theorem, differentiability condition is new. For iid, differentiability was guaranteed, but in general this may not be true.

<u>Remark</u>: GE theorem can be understood same way as the Cramér's Th.

<u>Remark</u>: Earlier examples of current / ~~tagged ptl~~ Many Brownian ptl is where Cramér's th do not apply (long-range correlation in time), but GE applies and gives ldf.

Remark: There are (Many) examples where GE does not work, and still there is ldf.

In such cases one has to calculate using inovative ways.

~~Integrability methods~~ [Betheansatz, Operator algebra, Techniques of integrability, Field theory, Novel numerical methods]

~~An~~ toy example where GE does not apply: (leave as excercise)

(but there is ldf)

$$P(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad P(y) = \frac{1}{2}\delta_{y,-1} + \frac{1}{2}\delta_{y,1}$$

$$M_n = \sum_i x_i + y \cdot n$$

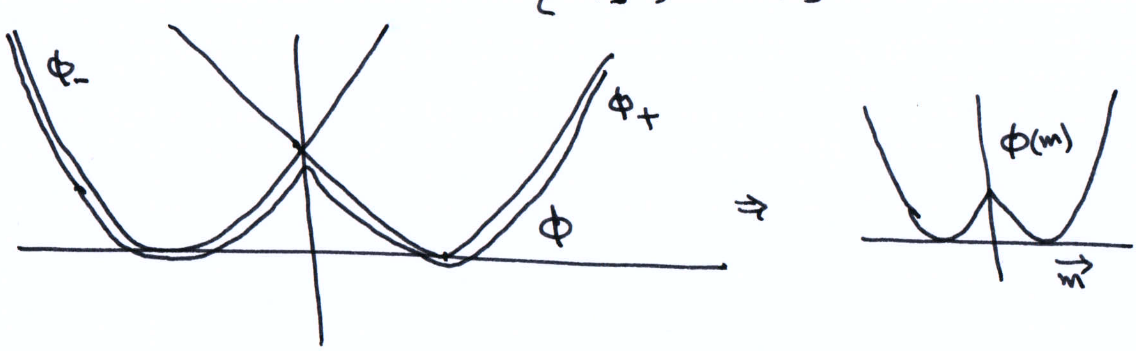
Show that

$$P\left(\frac{M_n}{n} = m \,\middle|\, y = \pm 1\right) \asymp e^{-n\phi_\pm(m)}$$

with $\phi_\pm(m) = \frac{1}{2}(m \pm 1)^2$

Then, $P(m) = \frac{1}{2} P(m|-1) + \frac{1}{2} P(m|+1) \asymp e^{-n\phi(m)}$

with $\phi(m) = \min\{\phi_+(m), \phi_-(m)\}$

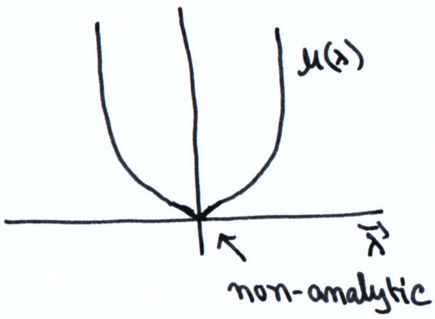


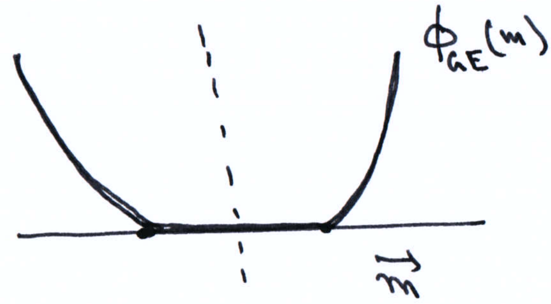
result using GE theorem.

$$\langle e^{\lambda M}\rangle = \frac{1}{2}\cdot e^{-\lambda n}\langle e^{\lambda x}\rangle^{n} + \frac{1}{2}e^{\lambda n}\langle e^{\lambda x}\rangle^{n}$$

$$= \cosh(\lambda n)\cdot e^{\lambda^2 \cdot \frac{n}{2}}$$

$$\Rightarrow \mu(\lambda) = \lim_{n\to\infty}\frac{\langle e^{\lambda M}\rangle}{n} = \frac{1}{2}\lambda^2 + |\lambda|$$

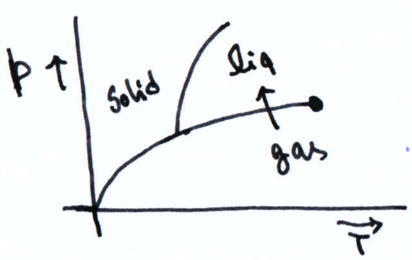




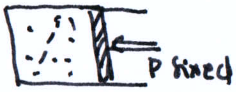
does not give the correct ldf.

Have you seen this before?

liq-gas transition [~~const~~ (P,T,N) ensemble]

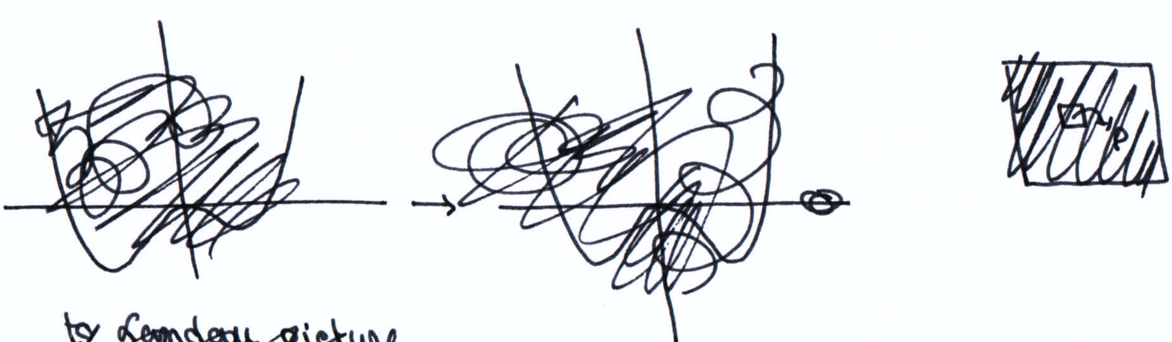


$\Rightarrow$



ρ is fluctuating

$P(\rho) \sim e^{-V\cdot\phi(\rho)}$

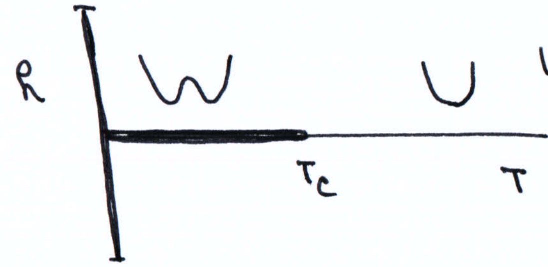


to Landau picture

~~Q: What happens if one takes coexistence into account?~~

~~[(V,T,N) ensemble]~~

$\boxed{\text{2d Ising model}}$



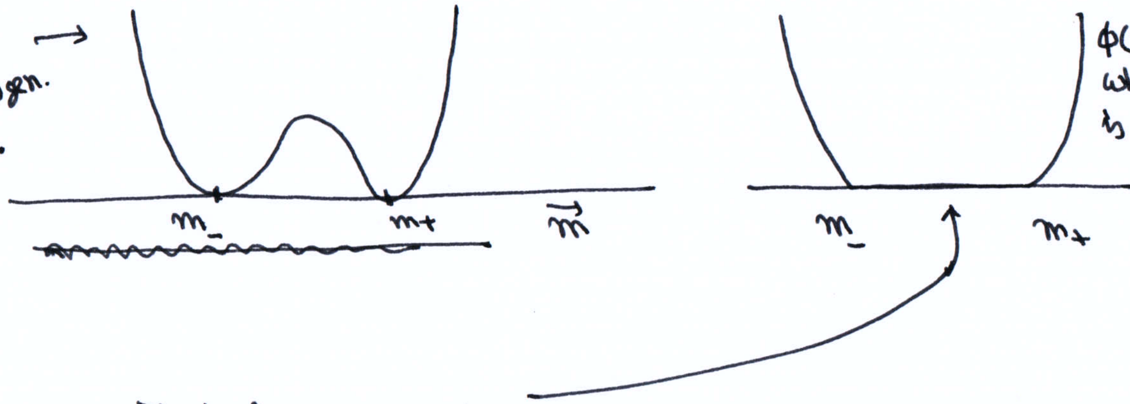
Landau free energy.

$(h=0, T)$ ensemble

$$\mathcal{P}\left(\frac{M}{V}=m\right) \asymp e^{-V\cdot\phi(m)}$$

Q. What happens if one includes coexistence.

$\phi(m)$ based on homogen. assumption.



$\phi(m)$ when coexistence is allowed.

It is less costly for system to ~~[illegible]~~ reach an intermediate value of $m_- < m < m_+$ by forming coexistence. This is because

$$\phi\left(\alpha m_- + (1-\alpha) m_+\right) > \alpha\,\phi(m_-) + (1-\alpha)\,\phi(m_+)$$

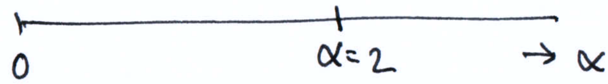
Remark: This is what one finds from exact solution of 2d Ising model. [Ref. see Fig 15 and Example 5.6 in Touchette review]

Remark: This gives an example where CLT does not work (non quadratic fluctuations near m^*), but it has ldf.

~~[illegible]~~ inverse of the case of $\frac{1}{x^{1+\alpha}}$, where CLT work but no ldf.

Summary chart

$$M_n = \sum_i x_i$$



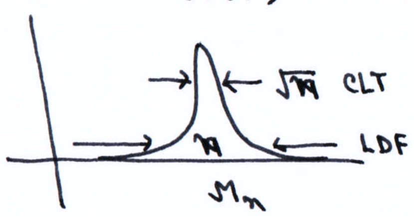
faster than power-law

$\frac{1}{x^{1+\alpha}}$ with $\alpha \geq 2$

$\frac{1}{x^{1+\alpha}}$ with $\alpha < 2$

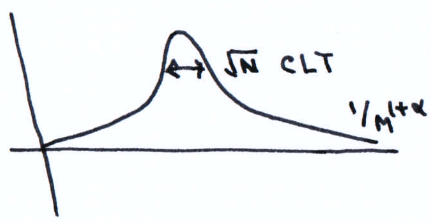
iid x

CLT works
LDF works



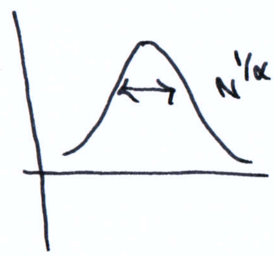
CLT works
No LDF

for large M_n
$P(M) \sim \frac{n}{M^{1+\alpha}}$



No CLT
No LDF

Lévy stable distributions.



Correlated x_i

CLT works if correlation short range.
(coarse-grain)

LDF works.

Strongly correlated
CLT may fail.
non-Gaussian fluctuations

but LDF may work.