9. Phase transitions and finite size scaling

One of the most common physical problems studied in simulations are *phase transitions* in various forms (ferromagnetism, Ising model, crystal melting, QCD . . .).

Most (but not all!)\(^1\) phase transitions can be described by an order parameter. Mathematically, this is zero in one phase (usually called the disordered phase), non-zero in the other phase (ordered phase). Thus, it cannot be an analytic function at the transition point. (Examples: magnetization in Ising model, Polyakov line in Ising gauge).

Normally, transitions are either 1st or 2nd order. The name comes from the number of derivatives of the free energy \( F = -T \log Z \) we need before we see discontinuous behaviour.

\(^1\)A common-day transition without an order parameter is the liquid-vapour 1st order phase transition, for example, boiling of water. There is no exact order parameter, and the two phases can be analytically connected. The transition line ends in a critical point, where the transition has 3-dim. Ising model universal behaviour.
• $F$ itself (zeroth derivative) is always continuous.

• First order — the order parameter (and almost any thermodynamical quantity) has a discontinuous jump:
  – latent heat: discontinuity in energy density

• Second order — second derivatives of $F$, i.e. various susceptibilities are divergent.
As a concrete example, the Ising model partition function with external field $H$ and $\beta = 1/T$

$$Z = \sum_{\{s_x=\pm 1\}} \exp \left[ -\beta \left( \frac{1}{2} \sum_{<xy>} (1 - s_x s_y) + H \sum_x s_x \right) \right]$$

gives magnetization $M$ and magnetic susceptibility $\chi_M$ as 1st and 2nd derivatives wrt. $H$:

$$M = \frac{1}{V} (dF/dH)_{H=0}$$

$$\chi_M = V (\langle M^2 \rangle - \langle M \rangle^2) = \frac{1}{V} (d^2F/dH^2)_{H=0}$$

(limit $V \to \infty$ implied here.)
Second order transitions are classified by their critical exponents, which characterize the behaviour at the critical point. The most important here are

- Magnetization: \( M \sim |T - T_c|^{\beta} \)
- Mag. susceptibility: \( \chi_M \sim |T - T_c|^{-\gamma} \)
- Heat capacity: \( C_V = \frac{1}{V} \frac{d\langle E\rangle}{dT} \sim |T - T_c|^{-\alpha} \)
- Correlation length: \( \xi \sim |T - T_c|^{-\nu} \)

For the 2d Ising model, these exponents are known exactly: \( \alpha = 0, \beta = 0.125, \gamma = 1.75, \nu = 1. \)

However, as already mentioned before, on a finite lattice we have finite number of degrees of freedom and everything is analytic! This causes several problems:

- What is a good order parameter? The order parameters are always either zero or non-zero. For example, both

\[
M = \left\langle \frac{1}{V} \sum_i s_i \right\rangle \\
|M| = \left\langle \left| \frac{1}{V} \sum_i s_i \right| \right\rangle
\]
are real order parameters in infinite volume, but $\langle M \rangle = 0$ and $\langle |M| \rangle > 0$ on a finite lattice. Of these, $\langle |M| \rangle$ is usable, since it is almost zero in the symmetric phase.

- How to locate the true phase transition?
- How to measure the critical exponents?
9.1. Finite size scaling in 2nd order phase transitions

For concreteness, let us consider Ising model. In infinite volume, the correlation length $\xi$ (domain size) diverges near the transition point as

$$\xi \propto |t|^{-\nu},$$

with $t = T - T_c \approx \beta_c - \beta$.

However, because the system in simulations has a finite size $L^d$, when the correlation length is $\xi \approx L$, the system already becomes effectively ordered. Thus, we can argue that the system has a pseudocritical point when

$$[\beta_c(\infty) - \beta_c(V)]^{-\nu} \propto L \implies \beta_c(V) = \beta_c(\infty) - \text{const.} \times L^{-1/\nu}$$

How to locate this point (if we don’t know $\nu$ or $\beta_c(\infty)$)? Consider, for example, magnetic susceptibility which diverges in infinite volume as

$$\chi_{|M|} = \frac{1}{V} \left( \langle M^2 \rangle - \langle |M| \rangle^2 \right) \propto |t|^{-\gamma}$$
We can now (somewhat arbitrarily) argue that on a finite volume $X_{|M|}$ has a maximum at the pseudocritical point $\beta_c(V)$. At this point the maximum value should be

$$\chi_{M,max} \propto (\beta_c(V) - \beta_c(\infty))^{-\gamma} \propto L^{\gamma/\nu}$$

The above (not extremely robust) argument gives us a prescription how to determine the true critical point $\beta_c = \beta_c(\infty)$ and even to estimate the critical exponents $\nu$ and $\gamma$:

1. Using various volumes $V$, locate the maximum of $X_{|M|}$.
2. Make a (power law) fit to the maximum location of $X_{|M|}$:

$$\beta_{\text{max}} = \beta_c - c_1 \times L^x$$

Fit has 3 parameters, $\beta_c$, $c_1$, $x$, where $x$ should be equal to $-1/\nu$.
3. The exponent $\gamma/\nu$ can be estimated from the maximum value $\chi_{\text{max}} \propto L^{\gamma/\nu}$.  

7
WARNING: while this process gives pretty good estimate of the infinite volume critical point $\beta_c$, the exponents can be more difficult to obtain reliably. This is especially so if we would use the heat capacity $C_V$ instead of $\chi_{|M|}$, this is due to the fact that the critical exponent $\alpha$ is usually much smaller than $\gamma$.

NOTE:
Above, somewhat misleadingly, we actually used susceptibility $\chi_{|M|} = \frac{1}{V} (\langle M^2 \rangle - \langle |M| \rangle^2 )$. This is strictly speaking not equivalent to the ’true’ magnetic susceptibility $\chi_M = \frac{1}{V} (\langle M^2 \rangle - \langle M \rangle^2 )$. In the broken phase these are equal, but in the symmetric phase these differ by a constant. The critical exponents are equal, however.
Example: 2d Ising model

2d Ising model with volumes $16^2$, $32^2$, $64^2$, $128^2$ and $256^2$. $\chi_M$ reweighted to a range of $\beta$-values around the critical point:

The peak of $\chi_M$ clearly grows (like $L_z$, note log-scale) and the location moves to larger $\beta$. 

![Graph showing the reweighted chi-magnetization for 16^2 to 256^2 Ising model]
Power law fit to the location of the maxima: \( \beta_{\text{max}} = \beta_c - c L^x \).
\((\chi^2 / \text{d.o.f} = 0.33 / 2, \text{ confidence level 0.84}.\)

Results:
\( \beta_c = 0.88093(24); \ x = -1.05(2). \)

Right results:
\( \beta_c = \ln(1 + \sqrt{2}) \approx 0.88137 \)
\( x = -1/\nu = -1. \)

We are very close, but still \( \sim 2\sigma \) off the correct values. This is most likely due to too small volume \((16^3)\).

If we drop \(16^3\) and fix the exponent \( x = -1 \), we obtain \( \beta_c = 0.88132(13) \), which is perfectly compatible with the right result.
Power law fit to the value of the maximum: $\chi_{\text{max}} = cL^z$.
If we exclude $16^2$, we obtain

$z = 1.740(8)$,
which is compatible with the right value

$z = \gamma/\nu = 1.75$

(Using also $16^2$ makes the fit a bit worse.)
Critical exponent $\beta$

Finite size scaling can be used to determine other exponents too: for example, let us consider the spin-spin correlation function just at the critical point:

$$\langle s_x s_y \rangle = G(|x - y|) \propto |x - y|^{-(d-2+n)}, \quad |x - y| \to \infty,$$

where $d$ is the dimensionality of the system and $\eta$ is the “anomalous” exponent of the correlation length (for 2d Ising, $\eta = 1/4(?)$).

We define the root mean square magnetization

$$M_{\text{rms}} = \sqrt{\langle M^2 \rangle} = \sqrt{\sum_{x,y} s_x s_y / V^2}.$$

At the infinite volume critical temperature $T_c$ (or $\beta_c$), we can approximate behaviour of the correlation function on a finite volume as

$$\sum_x \langle s_x s_y \rangle \propto \int_0^{L/2} dr \ r^{d-1} \ G(r) \propto \int_0^{L/2} dr \ r^{1-\eta} \propto L^{2-\eta}. $$
Thus, $M_{\text{rms}}$ becomes

$$M_{\text{rms}}^{T=T_c} \propto \sqrt{L^{2-d-\eta}} \propto L^{-\beta/\nu}.$$ 

In the last stage we used scaling law $2 - \eta = \gamma/\nu$ and the so-called *hyperscaling* $d\nu = 2\beta + \gamma$.

Hyperscaling is not always valid: this happens, for instance, in systems above their marginal dimensionality $d^*$ where the mean field values for the critical exponents become valid. For simple spin models with local action (Ising, for example) $d^* = 4$. 
Example: 2d Ising, again, and determine $M_{\text{rms}}$ from lattices $16^2 - 256^2$.

Power law fit: $M_{\text{rms}} = cL^z$.
The result from the fit is $z = -0.1255(4)$, which is compatible with the known value $z = -\beta/\nu = -0.125$

[Note: the plot is of log-log type]
**Binder cumulant**

Another frequently used method to determine the critical point is to use the intersection points of the *Binder cumulants*. This is often better than the maximum location of susceptibilities, because the finite size effects are usually much reduced.

The 4th order Binder cumulant $U_L$ is defined as

$$U_L = 1 - \frac{\langle s^4 \rangle}{3\langle s^2 \rangle^2}$$

(Here $s$ is the average spin; for Ising model $s = \frac{1}{L^d} \sum_x s_x$; for sigma models, $\bar{s} = \frac{1}{L^d} \sum_x \bar{\sigma}_x$, and $s^2 = \bar{s} \cdot \bar{s}$.)

$U_L$ behaves as follows:

- In the symmetric phase, $T > T_c$, $U_L = 0 + \mathcal{O}(1/V)$ as $L \to \infty$.
- In the broken phase, $T < T_c$, $U_L = 2/3 + \mathcal{O}(1/V)$ as $L \to \infty$.
- At the critical point, $U_L$ tends towards an universal value $0 < U_L^* < 2/3$. 
How to locate the critical point using $U_L$?

- Using various volumes $L^d$, calculate $U_L$’s as functions of $\beta$ (reweighting – remember that you have to reweight $\langle s^2 \rangle$ and $\langle s^4 \rangle$ separately!).

- Find the point where $U_L(\beta)$-curves cross. Usually, one finds the crossings using ascending pairs of volumes ($L_1/L_2$, $L_2/L_3$, . . . , where $L_1 < L_2 < L_3$ . . . ). These are extrapolated to $L \to \infty$, if needed. However, usually the finite volume dependence of the crossing point is very small!
**Example: 2d Ising model**

Left: $U_L$ measured from $16^2$ simulation. Right: volumes $16^2 - 256^2$, zooming in very close to the critical point $\beta_c$. Clearly, the slope of $U_L$ becomes *larger* as $L$ increases. The intersection points move only slightly.
How the intersection point evolves as $L$ increases? There is no clear finite $L$ systematics (within the statistical errors) at various $L$. The largest volume pair $128/256$ alone gives 
\[ \beta_c^{128/256} = 0.88144(6), \]
which is already practically compatible with the known value 
\[ \beta_c \approx 0.88137 \]
(Compare this to the maximum of $\chi |M|!$)
Justifying FSS: order parameter probability distributions

Where does the Binder cumulant behaviour come from? In general, we can approach the finite size scaling systematically using phenomenological order parameter probability distributions $P_L(s)$, which depend on the system size $L$ (or volume $V = L^d$).

- If $T > T_c$ and $L \gg \xi$, the correlation length, we are deep in the symmetric phase (remember that $\xi \propto |T - T_c|^{-\nu}$). We assume here $H = 0$. Now the probability distribution can be described as a Gaussian:

$$P_L(s) = \sqrt{\frac{V}{\pi \alpha}} \exp[-s^2V/\alpha]$$

$\alpha$ is a function of $\beta$ but independent of $L$, it is related to susceptibility:

$$\chi_M = V \int ds s^2 P_L(s) = \alpha/2.$$
If $T < T_c$ and $L \gg \xi$, we are deep in the broken phase. The probability distribution is a sum of 2 Gaussians, centered around magnetization $M_L$ near $M_{L=\infty}(\beta)$:

$$P_L(s) = \sqrt{\frac{V}{\pi \alpha}} \left[ \frac{1}{2} \exp[-(s - M_L)^2 V / \alpha] + \frac{1}{2} \exp[-(s + M_L)^2 V / \alpha] \right].$$

Region around $T \sim T_c$, $\xi \sim L$: the distribution cannot (necessarily) be described by simple Gaussians. The idea is that $P_L$, which is a function of $s, L, \xi$ ($T$ comes along through the temperature dependence of $\xi$), is a function of scaled variables $L/\xi, s \xi^{\beta/\nu}$:

$$P_L(s) = \xi^{\beta/\nu} g(L/\xi, s \xi^{\beta/\nu}) = L^{\beta/\nu} f(L/\xi, sL^{\beta/\nu}).$$

The power in front ensures that we can normalize $P_L$,

$$\int ds P_L(s) = 1,$$

and the last equality comes from the relation $L^{\beta/\nu} = (L/\xi)^{\beta/\nu} \xi^{\beta/\nu}$. 

20
From this we can immediately obtain the standard scaling relations

\[ \langle |s| \rangle_{T=T_c} \propto L^{-\beta/\nu} \]
\[ \chi_{|M|} = V(\langle s^2 \rangle - \langle |s| \rangle^2) \propto L^{\gamma/\nu} \]

From these distributions it is straightforward to obtain our results for the Binder cumulants:

- \( T > T_c \): \( U_L \rightarrow 0 \)
- \( T < T_c \): \( U_L \rightarrow 2/3 \)
- \( T \sim T_c \): \( U_L \approx 1 - \frac{\bar{\chi}_4(L/\xi)}{3(\bar{\chi}_2(L/\xi))^2} \)
- \( T = T_c \): \( U_L \rightarrow U^*_L = 1 - \frac{\bar{\chi}_4(0)}{3(\bar{\chi}_2(0))^2} \)

where \( \bar{\chi}_2 = L^{2\beta/\nu}\langle s^2 \rangle \), and \( \bar{\chi}_4 = L^{4\beta/\nu}\langle s^4 \rangle \).
How to get reliable error estimates in FSS?

In FSS, reweighting is presently the standard tool for locating maxima of susceptibilities/crossing points of some cumulants. Care should be taken in order to get reliable error estimates:

1. Divide the simulation data in jackknife or bootstrap blocks (see sect. 6!).
2. Reweight blocks separately to a suitable range of $\beta$-values.
3. Get the maximum value and location for each of the blocks.
4. Now the error estimate for the maximum location, using jackknife, for example, is

$$\delta \beta_{\text{max}} = \sqrt{\frac{N - 1}{N} \sum_n (\beta_{\text{max}}^{(m)} - \bar{\beta}_{\text{max}})^2}$$

where $\bar{\beta}_{\text{max}}$ is the average value of $\beta_{\text{max}}$. 

22
Locating the maximum in stage 3 can be cumbersome. The easiest way to do this is probably just to reweight using a very fine $\beta$-spacing, and using the maximum location. More accurate result can be reliably obtained by taking 3 points around the maximum and fitting a parabola to these points, and using the maximum location and value of the parabola. This is reliable because the reweighted curves are very smooth if the $\beta$-spacing is sufficiently fine.

Similar methods can be used for locating the crossing of cumulants. Now one locates the crossing point for a pair of volumes, jackknife/bootstrap block by block.
9.2. Finite size scaling in 1st order phase transitions

First order transitions are characterized by a discontinuity in the order parameter and thermodynamic densities, with an associated delta-peak behaviour in the susceptibility. Jump in energy density $\equiv$ latent heat.

However, at finite $V$ thermodynamic quantities (energy, say) become continuous and rounded. Instead of $\delta$-function behaviour in susceptibility ($C_V = d\langle E \rangle/dT$), there is just a 'hump'.

In simulations, this behaviour is visible only if the simulation time $\gg$ autocorrelation time $\tau$ at the transition point. This is typically very large, $\tau \propto \exp[-\sigma 2L^{d-1}]$, where $\sigma$ is the tension of the interface between the low-temperature and high-temperature phases.
**Probability distribution**

In first order transitions the correlation length remains finite in both the hot and cold phases. If $L \gg \xi$ in both phases, the probability distribution of the order parameter $P(s)$ at the critical point will be *mostly* a sum of 2 Gaussians, centered around “hot” and “cold” (symmetric and broken) expectation values.

However, if the order parameter is between the hot and cold phases, then the system is in a **mixed phase**:

At $T_c$ in the mixed phase, $P(s) \propto \exp[-\sigma 2L^{d-1}]$, where $\sigma$ is the interface tension and $2L^{d-1}$ is the total area of the interfaces (periodic boundary conditions!). $P(s)$ is $\sim$ flat, because the probability does not depend
on the fraction of the hot and cold phases (only through surface-surface interactions) and $s = (V_{\text{hot}}s_{\text{hot}} + V_{\text{cold}}s_{\text{cold}})/V$.

Note: this flat part becomes visible only if the volume is large enough. Usually the suppression of $P(s)$ in the middle is so strong that normal simulation methods are not good enough, and one has to use multi-canonical methods.

The behaviour of the probability distribution is actually used to measure the surface tension $\sigma$:

$$\sigma = \lim_{V \to \infty} \frac{1}{2L^{d-1}} \ln \frac{p_{\text{min}}}{p_{\text{max}}}.$$ 

Here $p_{\text{min}}$ is the minimum of the distribution between the peaks, and $p_{\text{max}}$ is the maximum. This formula has several subleading finite size corrections, which help to extrapolate the infinite volume limit [Bunk, Int.J.Mod.Phys.C3(1992)].
Potts model

Potts models are prototypes of models which exhibit a 1st order phase transition. A $q$-state Potts model is defined through

$$Z = \sum_{s_x=1\ldots q} \exp[-\beta E] \quad E = \sum_{<x,y>} [1 - \delta(s_x, s_y)]$$

Ising model is a Potts model with $q = 2$. In 2 dimensions, Potts models are self-dual, and the critical temperature is known: $\beta_c = \ln(1 + \sqrt{q})$. Also known are the correlation lengths and the latent heat. The transition is of second order if $q \leq 4$, otherwise first order. (In 3d the transition is second order only if $q = 2$.)

The transition is very “weakly first order” if $q$ is small. For example, when $q = 5$ the latent heat is $\mathcal{L} \approx 0.053$, interface tension $\sigma \approx 0.000199$, the correlation length in disordered phase $\xi \approx 2512$. The “natural” magnitude for all of these quantities is 1.

For $q = 20$, the transition is strong: $\mathcal{L} \approx 1.2$, $\sigma \approx 0.18$, $\xi \approx 2.7$. 

27
Probability distribution of the 20-state Potts model at the critical point \( \beta_c = \ln(1 + \sqrt{20}) \).

Clearly, the central part of the distribution approaches flat as the volume increases. However, the volumes needed are very large!

[A. Billoire, T. Neuhaus and B. Berg, 1994]
Probability distribution in 1st order transitions

As for the 2nd order phase transitions, we can construct phenomenological models for the probability distribution in 1st order transitions, which depend on the system size \( L (V = L^d) \).

- Let us look at a temperature-driven order-disorder transition. At the transition point (infinite \( V \)), \( \langle E \rangle \) jumps from \( E_- \) to \( E_+ \).

- In first order transitions \( \xi \) remains finite in both phases (unless there are Goldstone modes, which we neglect here). Let us assume that \( L \gg \xi \). Now we can describe the bulk phases with Gaussians

\[
P_+(E) \propto \exp \left[ -\frac{(E - \langle E(T) \rangle_+)^2 V}{2T^2C_+} \right]
\]

(equivalently for \( P_- \)). Here \( C_+ \) is the heat capacity of the hot bulk phase, \( C_+ = \partial \langle E \rangle / \partial T \). This implies that \( \langle E \rangle_+ = E_+ + C_+ \Delta T \).
Thus, neglecting the “surface contributions”, the probability near the critical temperature $T_c$ is a sum of 2 Gaussians:

$$P_L(E) \propto \frac{a_+}{\sqrt{C_+}} \exp \left[ -\frac{(E - E_+ - C_+ \Delta T)^2 V}{2T^2 C_+} \right]$$

$$+ \quad \frac{a_-}{\sqrt{C_-}} \exp \left[ -\frac{(E - E_- - C_- \Delta T)^2 V}{2T^2 C_-} \right]$$

Here $a_+$ and $a_-$ are the relative probabilities (weights) of the 2 phases. At $T = T_c$ the phases have equal probability. If we have $q_+$ degenerate symmetric phases and $q_-$ degenerate broken phases (for $q$-state Potts models, $q_+ = 1$, $q_- = q$), we have

$$a_+ = q_+ \exp \left[ -\frac{(E_+ - E_-) \Delta T V}{2T_c^2} \right]$$

$$a_- = q_- \exp \left[ -\frac{(E_+ - E_-) \Delta T V}{2T_c^2} \right]$$
Thus, we can calculate the expectation value of energy and $C$:

\[
\langle E \rangle_L = \frac{a_+ E_+ + a_- E_-}{a_+ + a_-} + \Delta T \frac{a_+ C_+ + a_- C_-}{a_+ + a_-}
\]

and

\[
C_L = \frac{\partial \langle E \rangle_L}{\partial T} = \frac{a_+ C_+ + a_- C_-}{a_+ + a_-} + \frac{q_+ q_- V}{T_c^2} \left[ E_+ - E_- + (C_+ - C_-) \Delta T \right]^2 \frac{1}{(a_+ + a_-)^2}
\]

The first term in $C_L$ just smoothly interpolates between the bulk values. The second term diverges as $V \to \infty$, it has a peak where $(a_+ - a_-)$ has a minimum value, or

\[
\Delta T \approx \frac{T_c^2 \ln(q_-/q_+)}{V} \frac{E_+ - E_-}{E_+ - E_-}
\]
with the height

\[ C^\text{max}_L \approx \frac{(E_+ - E_-)^2}{4T_c^2} V + \frac{C_+ - C_-}{2} \]

- Thus, in 1st order transitions, the peak of the susceptibility diverges as \( \propto V \), and the location behaves as \( 1/V \). Comparing to 2nd order critical exponents, this corresponds to \( \nu = d \) and \( \gamma = \nu \).

- NOTE: the behaviour of \( C^\text{max}_L \) is very easy to obtain from a straightforward scaling argument: at \( T_c \), the distribution \( P_L(E) \) is a sum of 2 Gaussians at \( E_+ \) and \( E_- \). Ignoring the shift \( \langle E \rangle \approx (E_+ + E_-)/2 \), and \( |E - \langle E \rangle| \approx (E_+ - E_-)/2 \). Thus,

\[ C_L \propto V \langle (E - \langle E \rangle)^2 \rangle \approx V \frac{1}{4}(E_+ - E_-)^2 \]
More precisely, there are several classes of mixed phases. In 3d, we have bubbles (droplets), cylinders and slabs. Here $\theta$ is the order parameter. The probability distribution is $P_X(\theta) = \exp[-\sigma A_X(\theta)]$, where $X$ is bubble, cylinder or slab, and $A_X$ is the area. For example, $A_{\text{bubble}}(\theta) = (4\pi)^{1/3}(3V)^{2/3}\theta^{2/3}$ where I normalized the bulk values $\theta_1 = 0$, $\theta_2 = 1$. 

\[
\log P(\theta) = -F(\theta)
\]
However, these geometric shapes are practically invisible in 'normal' volumes achieved in Monte Carlo simulations. Example of a bubble from a Monte Carlo simulation of a so-called cubic asymmetry model:
Identifying and locating first order transitions

Often it can be difficult to tell whether a transition is of first or second order (or if there is a transition at all!). Take, for example, $q = 5 \ 2$-dim. Potts model, where we know that the correlation length at the transition is $\sim 2500$. The lattice size must be much larger than this if we want to distinguish reliably the order of the transition!

Example: SU(2)-Higgs theory. This is an effective theory of the Standard Model of particle physics; it describes the high-temperature symmetry restoring phase transition of the SM. At $T < T_c \sim 100\text{GeV} \ (\sim 10^{12} \text{ K})$, the SM is in the broken phase (“Higgs condensate” $\leftrightarrow$ magnetization); at $T > T_c$ in the symmetric phase. In the early Universe this transition occurred $\sim 10^{-15}$ seconds after Big Bang.

The SU(2)-Higgs theory has been studied in detail 1994–2000 using lattice simulations. Since the Higgs particle has not yet been found, the phase diagram depends on the unknown mass of the Higgs $m_H$: 
When $m_H$ is small, the transition is “strong”, but gets weaker as $m_H$ increases. The 1st order transition line ends at a critical Higgs mass $m_{H, \text{crit}} \approx 72$ GeV. At larger Higgs masses there is no transition (cf. water-vapour transition).

Let us look at the point at $m_H \sim 52$ GeV (in the following plots this is labeled as 60 GeV, due to outdated historical reasons).
The probability distribution of the Higgs field squared $|\phi|^2$ ("magnetization") at $T_c$ has a clear 2-peak structure, which becomes stronger as the volume is increased:
The maximum of the susceptibility diverges as $1/V$ (in the figure $C(R^2) = \chi|\phi^2|/V$).
Locating the critical point: we can use the maximum location of the susceptibilities (here we have 2 order parameters, $R = |\phi|$ and $L$, and hence 2 susceptibilities). These behave as $1/V$.

However, a more precise method is to use *equal weight* histogram method: for each volume, find the temperature where the probabilities (area of the peaks) of the 2 phases are equal. This has only corrections $\propto \exp[-\sigma A]$. 
At $m_H \sim 68$ GeV (here labelled as 70 GeV) the transition is clearly weaker: the separation of the peaks is not nearly as strong as before.