

In statistical physics, we are often dealing with systems that comprise many components. In order to calculate their statistics, high-dimensional integrals over those variables $\mathbf{x} \in \mathbb{R}^N$ with $N \gg 1$ are required. A typical form is

$$\mathcal{Z} = \int_{\mathbf{x}} e^{S_0(\mathbf{x}) + S_V(\mathbf{x})},$$

where we introduced $\int_{\mathbf{x}} := \int_{-\infty}^{\infty} d\mathbf{x}$ for brevity.

In general, $S_V(\mathbf{x})$ couples variables in such a way that the integration does not factorize. If N is small, such integrals would be tractable by numerical methods still, so there no real problem exists. This ceases to be the case in higher dimensions, where the curse of dimensionality manifests. Luckily, there are approximation methods for this case. Mean-field theory is a particular example for this, which will be the subject of this post.

To illustrate how DMFT can be of help here, let's first consider a toy problem where we just have a single coupled $S_V(\mathbf{x})$ and set $S_0(\mathbf{x}) = 0$, with

$$S_V(\mathbf{x}) = -\mathbf{x}^T A \mathbf{x} = -x_i A_{ij} x_j$$

where we set $A_{ij} = \delta_{ij}$ for simplicity. This introduces an interaction between variables x_i and x_j that prevents factorization and the integral from being tractable, i.e. the associated probability density $p(\mathbf{x}) \propto e^{-x_i A_{ij} x_j}$ does not factorize. (Note that in the case of a positive definite A the integral is a Gaussian and exactly solvable, but we will still use the approximation for illustration.)

Decoupling

To mitigate this, DMFT introduces an **auxiliary field** that caps the threads between x_1 and x_2 . To accomplish this, we use the identity

$$\delta(\mathbf{y} - \mathbf{x}) = \int_{\tilde{\mathbf{y}}} \exp(-i\tilde{\mathbf{y}}^T (\mathbf{y} - \mathbf{x})).$$

Further down, we will attempt a more intuitive picture for the identity.

The core procedure goes in two steps: **decoupling** (line 2) and **isolation** (line 4).

$$\begin{aligned}
\mathcal{Z} &= \int_{\mathbf{x}} e^{-\mathbf{x}^T A \mathbf{x}} \\
&= \int_{\mathbf{x} \mathbf{y}} e^{\mathbf{y}^T \mathbf{y}} \delta(\mathbf{y} - A \mathbf{x}) \\
&= \int_{\mathbf{x} \mathbf{y} \tilde{\mathbf{y}}} e^{\mathbf{y} - i \tilde{\mathbf{y}}^T (\mathbf{y} - A \mathbf{x})} \\
&= \int_{\mathbf{y} \tilde{\mathbf{y}}} e^{-i \tilde{\mathbf{y}}^T \mathbf{y} + \mathbf{y}^T \mathbf{y}} \int_{\mathbf{x}} e^{i \tilde{\mathbf{y}}^T A \mathbf{x}} \\
&= \prod_i^N \int_{y_i \tilde{y}_i} e^{-i \tilde{y}_i y_i} \int_{x_i} e^{y_i y_i + i \tilde{y}_i x_i} \\
&= \left(\int_{y \tilde{y}} e^{-i \tilde{y} y + y y} \int_x e^{y y + i \tilde{y} x} \right)^N \\
&= \int_{y \tilde{y}} \underbrace{e^{-N i \tilde{y} y} e^{N \ln Z_1(y, \tilde{y})}}_{=: I(y, \tilde{y})}.
\end{aligned}$$

In the last step, we have identified the **single-site partition function** Z_1 , which apparently goes in as an ingredient to the overall partition function \mathcal{Z} . The auxiliary field \tilde{y} describes how those subsystems are coupled to one another.

So far, all manipulations were exact. Unfortunately, to make progress, we now have to introduce an approximation. This is called the **saddle point approximation**, and approximates the integral with the point where the integrand is stationary,

$$\begin{aligned}
\partial_y I(y^*, \tilde{y}^*) &= 0 \\
\partial_{\tilde{y}} I(y^*, \tilde{y}^*) &= 0.
\end{aligned}$$

Naively, this derives from the intuition that the integrand should be maximal. However, the integral in \tilde{y} is complex, so what is the meaning of the variation with respect to \tilde{y} ? Looking more closely, this is where the name **stationary phase approximation** becomes more appropriate. Essentially, this picks out the value where the integrand oscillates most slowly, and its values don't average out.

In any case, this gives a set of **saddle point equations of the form**

$$\begin{aligned}
\tilde{y}^* &= \frac{1}{Z_1(y, y^*)} \int_x \frac{1}{2} y^* e^{S_0(x) + y^* y^* + i \tilde{y}^* x} = \frac{1}{2} y^* \langle 1 \rangle_{S_1} = \frac{1}{2} y^*. \\
y^* &= \frac{1}{Z_1(y, y^*)} \int_x x e^{S_0(x) + y^* y^* + i \tilde{y}^* x} = \langle x \rangle_{S_1},
\end{aligned}$$

where we defined $S_1(x; y, \tilde{y}) = S_0(x) + y^* y^* + i \tilde{y}^* x$.

Note that curiously, y^* is the average value of x under this effective, one-dimensional theory S_1 . This reflects that we have decoupled the variables, so that each variable sees an independent distribution S_1 , summarizing the *mean-field* influence of the others.

Relation to conditional independence One can point out a relationship to conditional independence here: If we consider a probability $p(x_i, x_j) \propto \exp(-x_i A_{ij} x_j)$, the variables are said to be statistically dependent, as the distribution does not factorize. In the context of graphical models, such dependence typically comes about from an unobserved latent confounder, coupling the variables. From the transform above, we can identify \mathbf{y} as the confounder that couples the variables through $\tilde{\mathbf{y}}$. Conversely, fixing or observing a value of $(\mathbf{y}, \tilde{\mathbf{y}})$ in the integral makes x_i and x_j conditionally independent, $p(x_1, x_2 | \mathbf{y}, \tilde{\mathbf{y}}) = p(x_1 | \mathbf{y}, \tilde{\mathbf{y}}) p(x_2 | \mathbf{y}, \tilde{\mathbf{y}})$.

After this brief introduction, let's take a look how we can apply this to an actual physical problem.

Example: Ising model

We can apply the developed formalism to the Ising model. It can be described again in terms of a partition function that depends on a parameter \mathbf{j} ,

$$\mathcal{Z}(\mathbf{j}) = \int_{\mathbf{x}} e^{-\mathbf{x}^T J \mathbf{x} + \mathbf{j}^T \mathbf{x}}.$$

Here, the term $\mathbf{x}^T J \mathbf{x} = x_i J_{ij} x_j$ takes the role of the coupling A between variables we have seen before.

To illustrate the main difficulty and how the mean-field approach can solve it, we will first discuss the ad-hoc approach that is often introduced in undergraduate courses. Afterwards, we will see how this relates to the formal approach developed above.

Ad-hoc mean-field theory

The mean-field approach typically gets introduced in an ad-hoc fashion, in the sense that one replaces the value of each spin in the partition function by its average $\bar{\mathbf{x}}$ plus a small fluctuation $\delta \mathbf{x}$, merely amounting to a shift of integration variables. Subsequently, one discards quartic terms in $\delta \mathbf{x}$, so that we get

$$\mathcal{Z}(\mathbf{j}) = \int_{\bar{\mathbf{x}} + \delta \mathbf{x}} e^{-\bar{\mathbf{x}}^T J \bar{\mathbf{x}} - 2\bar{\mathbf{x}}^T J (\bar{\mathbf{x}} + \delta \mathbf{x}) + \mathbf{j}^T (\bar{\mathbf{x}} + \delta \mathbf{x})}.$$

This renders the partition function tractable, and we obtain the expectation by the self-consistency condition

$$\bar{\mathbf{x}} \equiv \langle \bar{\mathbf{x}} + \delta \mathbf{x} \rangle_{\delta \mathbf{x}} = \partial_{\mathbf{j}} \ln \mathcal{Z}(\mathbf{j})|_{\mathbf{j}=0},$$

yielding

$$\begin{aligned} \bar{\mathbf{x}} &= \frac{1}{\mathcal{Z}(0)} \int_{\bar{\mathbf{x}} + \delta \mathbf{x}} e^{-2\bar{\mathbf{x}}^T J (\bar{\mathbf{x}} + \delta \mathbf{x})} \\ &= \tanh(-2J\bar{\mathbf{x}}). \end{aligned}$$

This exposes $-2J\bar{\mathbf{x}}$ as an **effective mean field** that controls the statistics of $\bar{\mathbf{x}} + \delta \mathbf{x}$.

A more formal approach

We can also solve the Ising model by the formal technique developed above. To this end, we again first decouple the action and then identify a field that influences all sites.

We start again from

$$\begin{aligned}
\mathcal{Z}(\mathbf{j}) &= \int_{\mathbf{x}} e^{-\mathbf{x}^T J \mathbf{x} + \mathbf{j}^T \mathbf{x}} \\
&= \int_{\mathbf{y}, \tilde{\mathbf{y}}} \int_{\mathbf{x}} e^{-\mathbf{x}^T \mathbf{y} - i \tilde{\mathbf{y}}^T (\mathbf{y} - J \mathbf{x}) + \mathbf{j}^T \mathbf{x}} \\
&= \int_{\mathbf{y}, \tilde{\mathbf{y}}} e^{-i \tilde{\mathbf{y}}^T \mathbf{y}} \int_{\mathbf{x}} e^{-\mathbf{x}^T \mathbf{y} + i \tilde{\mathbf{y}}^T J \mathbf{x} + \mathbf{j}^T \mathbf{x}}.
\end{aligned} \tag{1}$$

Again, we decoupled **via** $\mathbf{y} := J \mathbf{x}$ in line 2. Unfortunately, we cannot yet decouple the later integrand, as $J \mathbf{x} = J_{ij} x_j$ mixes the variables. To move forward, we have introduce one more trick:

Averaging over the quenched disorder

If a problem is too hard, it makes sense to try solving an approximate version of it. In statistical physics, this usually means trading exact results for a statistical answer, in other words: **ignoring microscopical details of the system**. Practically, this means that we say that we only know J up to its statistics, and hence are interested in answers **in expectation** over them.

It is common to assuming that its entries are identically, independently, and Gaussian distributed,

$$J_{ij} \sim \mathcal{N}(0, \sigma^2/N),$$

an assumption that is backed partially by the central limit theorem and partially by reasons of analytical convenience.

With this, we can approach the J -dependent term by analytically solving a multivariate Gaussian integral

$$\begin{aligned}
&\langle e^{i \tilde{\mathbf{y}}^T J \mathbf{x}} \rangle_{J_{ij} \sim \mathcal{N}(0, \sigma^2/N)} \\
&= \langle e^{i \tilde{\mathbf{y}}^T \mathbf{J}} \rangle_{J_i \sim \mathcal{N}(0, x_i^2 \sigma^2/N)} \\
&= e^{-\frac{1}{2} \sigma^2 / N x_i^2 \tilde{y}_i^2},
\end{aligned}$$

which is called the **Hubbard-Stratonovich trick**.

This finally allows us to decouple

$$\begin{aligned}
\langle \mathcal{Z}(\mathbf{j}) \rangle_J &= \int_{\mathbf{y}\tilde{\mathbf{y}}} e^{-i\tilde{\mathbf{y}}^T \mathbf{y}} \int_{\mathbf{x}} e^{-\mathbf{x}^T \mathbf{y} + i\tilde{\mathbf{y}}^T J\mathbf{x} + \mathbf{j}^T \mathbf{x}} \\
&= \int_{\mathbf{y}\tilde{\mathbf{y}}} e^{-i\tilde{\mathbf{y}}^T \mathbf{y}} \prod_i^N \int dx_i e^{-x_i y_i - \tilde{y}_i^2 \sigma^2 / N x_i^2 + j_i x_i} \\
&= \int_{\mathbf{y}\tilde{\mathbf{y}}} e^{-N i \tilde{\mathbf{y}}^T \mathbf{y}} e^{N \ln Z_1(\mathbf{y}, \tilde{\mathbf{y}})}.
\end{aligned}$$

In the last line, we have again identified the single site partition function. Finally, we perform the saddle point approximation. After taking the derivative, we discard the subleading term $\propto \sigma^2/N$, and get

$$\begin{aligned}
\tilde{y}^* &= \frac{1}{Z_1(y^*, \tilde{y}^*)} \int_x -x e^{-xy^* + jx} = \langle -x \rangle_{S_1} \\
y^* &= \frac{1}{Z_1(y^*, \tilde{y}^*)} \int_x -2\tilde{y}\sigma^2/N x^2 e^{-xy^* + jx} \\
&= \frac{1}{Z_1(y^*, \tilde{y}^*)} \int_x 2\langle x \rangle_{S_1} (Jx)^2 e^{-xy^* + jx} \\
&= 2\langle x \rangle_{S_1} \sigma^2 / N \langle x^2 \rangle_{S_1}.
\end{aligned}$$

Recalling the introduction of \mathbf{y} as $\mathbf{y} = J\mathbf{x}$, we can make sense of this

$$y_i = J_{ij}x_j = J_{ij}(J_{jk}x_k J_{kl}x_l) = \sigma^2 x^2$$

Equivalently, we can write the last equation as

$$y^* = \cosh^{-2}(\sigma^2 y^*).$$

Plugging in this solution into the integral, we end up with

$$\begin{aligned}
\langle \mathcal{Z}(j) \rangle_{J, \text{MF}} &= \left(\int_x e^{-xy^* + jx} \right)^N \\
&= (2 \cosh(-y^* + j))^N.
\end{aligned}$$

From this, we can see that y^* naturally appears as a field influencing the statistics of x . In particular, we get

$$\bar{x} = \partial_j \ln \langle \mathcal{Z}(j) \rangle_{J, \text{MF}}|_{j=0} = \tanh(-y^*).$$

Advanced MFT cookbook

When tackling more complicated problems, there are a couple of identities that can take a large part of the way in decoupling an integral. These are

$e^{-\frac{1}{2}y\sigma^2y} = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{\tilde{y}} e^{-\frac{1}{2}\frac{1}{\sigma^2}\tilde{y}^2 + i\tilde{y}y}$ **decoupling (Hubbard-Stratonovich transform)**

$$\langle e^{iy^T \tilde{y}} \rangle_{\tilde{y} \sim \mathcal{N}(0, A)} := \frac{1}{(2\pi \det A)^{d/2}} \int d\tilde{y} e^{-\frac{1}{2}\tilde{y}^T A^{-1} \tilde{y} + iy^T \tilde{y}} \equiv e^{-\frac{1}{2}y^T A y}.$$

$$\langle e^{iy^T \tilde{y}} \rangle_{\tilde{y} \sim \mathcal{N}(0, A)} = e^{-\frac{1}{2}y^T A y}$$

(Note that this can be seen from the general definition of **cumulant generating functions** $W(j)$)

$$\langle e^{jx} \rangle = Z(j) =: e^{W(j)},$$

which for $y \sim \mathcal{N}(0, A)$ precisely gives

$$\langle e^{jx} \rangle = \frac{1}{\sqrt{|2\pi\Sigma|}} \int_x e^{-\frac{1}{2}xA^{-1}x} e^{jx} = e^{\frac{1}{2}j\Sigma j}.$$

Note that this in this case is **just the Fourier transform of the Gaussian**, where we know that the variance inverts.

This trick can be especially useful to get rid of the inverse A^{-1} at the cost of introducing an integral. To remember the workings of this formula, note that the integrating exchanges x for y in the Gaussian, removes the linearity and inverts the variance.)

As a special case, we can take the denominator to the other side and get with $A \rightarrow 0$ and get

$$\frac{1}{\sqrt{|2\pi A|}} e^{-\frac{1}{2}y^T A^{-1}y} \rightarrow \delta(y) = \int d\tilde{y} e^{iy^T \tilde{y}},$$

recovering the characterization of the δ -function.

$$\delta(y) = \int_{\tilde{y}} e^{i\tilde{y}y} \text{ **decoupling**}$$

The Hubbard-Stratonovich transform is useful to decouple quadratically coupled variables, or to trade a matrix inversion for an integral. However, not all interactions $S_V(x) = y(x)$ can be solved this easily. Using the δ -identity, we can still formally decouple the integral:

$$\begin{aligned} & \int dx e^{S_0(x) + y(x)} \\ &= \int dx \int dy e^{S_0(x) + y} \delta(y - y(x)) \\ &= \int dx \int dy \int d\tilde{y} e^{S_0(x) + i\tilde{y}(y - y(x))} \\ &= \int dy \int d\tilde{y} e^{i\tilde{y}y} C(\tilde{y}). \end{aligned}$$

Using this, it is for example possible to derive the Fourier transform of a Gaussian, using the constraint $y \stackrel{!}{=} y(x) = A^{-1}x$, thus recovering the Hubbard-Stratonovich identity,

$$\begin{aligned}
e^{-\frac{1}{2}x^T A^{-1}x} &= \int dy e^{-\frac{1}{2}y^T A y} \delta(y - A^{-1}x) \\
&= \int dy d\tilde{y} e^{-\frac{1}{2}y^T A y + i\tilde{y}(y - A^{-1}x)} \\
&= \int d\tilde{y} e^{-\frac{1}{2}\tilde{y}^T A^{-1}\tilde{y} - i\tilde{y}A^{-1}x} \\
\tilde{y} \rightarrow A\tilde{y} &= \int d\tilde{y} e^{-\frac{1}{2}\tilde{y}^T A\tilde{y} - i\tilde{y}x},
\end{aligned}$$

where we did a transformation of variables in the last step.

Finally, it is interesting to see how these two approaches relate: Any hard δ -constraint may be implemented as a smooth “relaxation” via a Hubbard-Stratonovich transform:

$$\begin{aligned}
&\int dx \int dy e^{S_0(x) + y(x)} \delta_\sigma(y - y(x)) \\
&= \int dx \int dy e^{-\frac{1}{2\sigma^2}(y - y(x))^2} \\
&\stackrel{\text{H.S.}}{=} \int dx \int dy \int d\tilde{y} e^{-\frac{1}{2}\sigma^2\tilde{y}^2 + i\tilde{y}(y - y(x))} \\
&= \int dy \int d\tilde{y} e^{-\frac{1}{2}\sigma^2\tilde{y}^2 + i\tilde{y}y} C(\tilde{y}).
\end{aligned}$$

The first term becomes approximately constant as we take the limit $\sigma \rightarrow 0$, approximating the δ -function.

Relation to Lagrange multipliers

Suppose we strive to optimize a function $\mathcal{L}(x)$ under a constraint $y(x) = 0$. The method of Lagrange multipliers solves this problem by introducing dummy parameter \tilde{y} . Then, an augmented objective is to be minimized

$$\tilde{\mathcal{L}}(x; \tilde{y}) = \mathcal{L}(x) + \tilde{y}y(x)$$

by requiring $\partial_x \tilde{\mathcal{L}}(x^*; \tilde{y}^*) = 0$ and $\partial_{\tilde{y}} \tilde{\mathcal{L}}(x^*; \tilde{y}^*) = 0$. The second condition by construction amounts to fulfilling the constraint. There is however a caveat that we will dig into below.

How does this relate to our former method? Let us rewrite the objective as an integral by superimposing all possible solutions,

$$\int dx e^{-\mathcal{L}(x)} \delta(y(x) - 0).$$

We can apparently reframe the problem as finding the x that maximizes the integrand. I (It will become clear in a moment while it is useful to exponentiate the loss.)To this end, let's again raise the δ -constraint,

$$\int dx \int d\tilde{y} e^{-\mathcal{L}(x) + i\tilde{y}(y(x) - 0)} =: \int dx \int d\tilde{y} e^{-\tilde{\mathcal{L}}(x)}.$$

At this point, we can look for the point that maximizes the integrand. This leads to saddle-point equations

$$\begin{aligned}\partial_x \tilde{\mathcal{L}}(x^*; \tilde{y}^*) &= 0 \\ \partial_{\tilde{y}} \tilde{\mathcal{L}}(x^*; \tilde{y}^*) &= 0.\end{aligned}$$

This is precisely the prescription of the Lagrange paradigm! The Lagrange multiplier can hence be thought of as an auxiliary field in disguise that penalizes the objective's values to comply with the constraint. Crucially however, we have no guarantee that the saddle point equations have a solution.

This perhaps also reveals that an optimization via Lagrange multipliers may not have a solution. What to do then? We could employ a δ_σ -relaxation, which will ensure that an optimum exists. This allows for slightly (according to σ) violating the constraint until a minimum exists.