# Statistical Physics of Computation - Exercises

Emanuele Troiani, Vittorio Erba

September 2024

## Week 1

### Introduction to the saddle point method

#### Basic idea

This exercise will introduce a very useful tool to compute asymptotics of integral and show it in practice in an example. Suppose we want to compute the integral $I_\beta$ for $\beta \gg 1$:

$$I_\beta = \int_{\mathbb{R}} e^{-\beta f(t)}\, dt$$

for a reasonably regular function $f(t)$ on $\mathbb{R}$.

1. Intuitively, what values of $f(t)$ will affect the integral the most?

   The values that minimize $f(x)$.

Call $T_0 = \arg\min_t f(t) \subset \mathbb{R}$ the set of points for which $f(t)$ is (globally) minimized. If $f(t)$ is not bounded from below, i.e. $T_0 = \emptyset$, the integral is infinite. We assume $T_0 = \{t_0\}$, that is there is a unique global minimum.

2. Taylor expand $f(t)$ around $t_0$. Argue that if $f''(t_0) > 0$, then

   $$I_\beta \approx e^{-\beta f(t_0)} \int_{\mathbb{R}} e^{-\beta f''(t_0) t^2/2}\, dt$$

   where the corrections to this integral are exponentially small in $\beta$.

   Write:

   $$f(t) = f(t_0) + \frac{1}{2} f''(t_0) t^2 + \mathcal{O}(t^2)$$

   where we used that $f'(t_0) = 0$ because $f$ is stationary in $t_0$. We obtain the result by truncating this expression at second order and plugging it into the integral.
   Let's now take a step back and reflect on what we just did. If we were to be careful there are the following steps:

   (a) Split the integral over $\mathbb{R}$ in $I_0$, an interval around $t_0$, and $\mathbb{R} \setminus I_0$

   (b) Realise that the integral over $\mathbb{R} \setminus I_0$ is asymptotically vanishing

   (c) Realise that the integral over $I_0$ of the original function is the same as the one of the function expanded to the second order

(d) Realise that the integral of the approximated function on $I_0$ is the same as the one on $\mathbb{R}$

3. Conclude that

$$I_\beta \approx \sqrt{\frac{2\pi}{\beta f''(t_0)}} e^{-\beta f(t_0)}$$

Recall that $f''(t_0) > 0$, so the integral is a simple Gaussian integral. The result follows from the well known formula valid for $a > 0$:

$$\int_{\mathbb{R}} e^{-at^2}\, dt = \sqrt{\frac{\pi}{a}}$$

4. Suppose $T_0 = \{t_0, t_1\}$ with $t_0 \neq t_1$. As a consequence of the previous question, why do we have that?

$$I_\beta \approx \sqrt{\frac{2\pi}{\beta f''(t_0)}} e^{-\beta f(t_0)} + \sqrt{\frac{2\pi}{\beta f''(t_1)}} e^{-\beta f(t_1)}$$

Since exponential functions are rapidly decreasing we can consider any integral around the minimum as integrals over the whole real line. As a consequence we can use the Gaussian integration formula once again.

**Concentration though the saddle point**

In the class we will typically study systems with characteristic size $N \gg 1$, and study quantities of the form $\langle f(x) \rangle$:

$$\langle f(x) \rangle = \frac{\int \mathrm{d}x\, f(x) e^{N\phi(x)}}{\int \mathrm{d}x\, e^{N\phi(x)}} \tag{1}$$

1. Show that if $N$ is large enough, then $\langle f(x) \rangle = f(x_0)$, where $x_0$ is the global maximum of $\phi(x)$ As the name of the exercise suggests, we want to use the saddle point. We can compute the denominator, which will simply be

$$\int \mathrm{d}x\, e^{N\phi(x)} = C e^{N\phi(x_0)} \tag{2}$$

where $C$ doesn't really matter in this exercise. For the numerator, adding $f(x)$ to the integral doesn't change the saddle point as it's independent of $N$. In fact:

$$\int \mathrm{d}x\, f(x) e^{N\phi(x)} = \int \mathrm{d}x\, e^{N[\phi(x) + \log f(x)/N]} \tag{3}$$

So the saddle point is still $x_0$, and

$$\int \mathrm{d}x\, f(x) e^{N\phi(x)} = C f(x_0) e^{N\phi(x_0)} \tag{4}$$

This gives the result.

2. What would happen if $\phi(x)$ has two global maxima $\{x_1, x_2\}$?

   In this case we can use point 4 of the previous exercise. Notice that having two global maxima means $\phi(x_1) = \phi(x_2)$, but the curvature near these points can look quite different. For the denominator we have

$$\sqrt{\frac{2\pi}{N}} e^{N\phi(x_1)} \left( \frac{1}{\sqrt{\det \nabla^2 \phi(x_1)}} + \frac{1}{\sqrt{\det \nabla^2 \phi(x_2)}} \right) \tag{5}$$

   while for the numerator

$$\sqrt{\frac{2\pi}{N}} e^{N\phi(x_1)} \left( \frac{f(x_1)}{\sqrt{\det \nabla^2 \phi(x_1)}} + \frac{f(x_2)}{\sqrt{\det \nabla^2 \phi(x_2)}} \right) \tag{6}$$

   A compact way to write the result is to introduce the ratio of the curvatures $\gamma$

$$\gamma = \sqrt{\frac{\det \nabla^2 \phi(x_1)}{\det \nabla^2 \phi(x_2)}} \tag{7}$$

   then

$$\langle f(x) \rangle = \frac{f(x_1) + f(x_2)\gamma}{1 + \gamma} \tag{8}$$

**Stirling's formula**

Let's use the saddle point method to derive a famous approximation of the factorial.

1. Show that for $n \in \mathbb{N}$, $n! = \int_0^\infty x^n e^{-x}\, dx$

   We do it by induction: first notice that:

$$0! = \int_0^\infty e^{-x}\, dx = e^0 = 1$$

   then with integration by parts :

$$n!(n+1) = (n+1) \int_0^\infty x^n e^{-x}\, dx = \int_0^\infty x^{n+1} e^{-x}\, dx = (n+1)!$$

2. Write $n! = n^{n+1} \int_0^\infty e^{-nf(x)}\, dx$ for a certain function $f(x)$

   We first do some manipulations:

$$\int_0^\infty x^n e^{-x}\, dx = \int_0^\infty e^{n \log x - x}\, dx$$

   Now we want all the terms in the exponent to scale with the same power of $n$, so we do a change of variable $x \to nx$:

$$n \int_0^\infty e^{n \log x - nx + n \log n}\, dx = n^{n+1} \int_0^\infty e^{-nf(x)}\, dx$$

   where $f(x) = x - \log x$

3. Use the saddle point method to show that for $n \gg 1$ we have:

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

Let's study $f(t)$:

$$f'(t) = 1 - \frac{1}{x}$$

$$f''(t) = \frac{1}{x^2}$$

$$f'(t) = 0 \implies x = 1$$

$$f(1) = f''(1) = 1$$

So with using the saddle point method we get:

$$n! \approx \sqrt{\frac{2\pi}{n}} n^{n+1} e^{-n}$$

We get the final formula using some algebraic manipulation

## Entropy and free entropy

In this exercise, we review some useful relationship between entropy and free entropy. Recall that, given a system with degrees of freedom $s$ and Hamiltonian $\mathcal{H}[s]$, the free entropy is defined as

$$\Phi = \log \mathcal{Z} = \log \int ds\, e^{-\beta\mathcal{H}[s]}\,, \tag{9}$$

where we also defined the partition function $\mathcal{Z}$. Recall that the Hamiltonian is extensive in the thermodynamic limit, i.e. $\mathcal{H}[s] = \mathcal{O}(N)$.

1. Show that for any model with free entropy $\Phi$ we have:

$$\langle \mathcal{H} \rangle = -\frac{\partial \Phi}{\partial \beta}\,, \tag{10}$$

where the angular average is w.r.t. the Gibbs distribution

$$\langle f \rangle = \frac{\int ds\, e^{-\beta\mathcal{H}[s]} f(s)}{\int ds\, e^{-\beta\mathcal{H}[s]}}\,. \tag{11}$$

Is this relationship true for all $N$, or only in the thermodynamic limit $N \to \infty$?

We have the following identity:

$$\frac{\partial}{\partial \beta}\mathcal{Z} = \int \frac{\partial}{\partial \beta} e^{-\beta\mathcal{H}[s]}\, ds = -\int \mathcal{H}[s] e^{-\beta\mathcal{H}[s]}\, ds = -\mathcal{Z}\langle E \rangle$$

So we have:

$$\langle \mathcal{H} \rangle = -\frac{1}{\mathcal{Z}}\frac{\partial \mathcal{Z}}{\partial \beta} = -\frac{\partial \log \mathcal{Z}}{\partial \beta} = -\frac{\partial \Phi}{\partial \beta}$$

This is true for any $N$.

2. Defining the entropy at fixed energy $S(E)$ as the logarithm of the number of configurations at energy $E$, show that you can write the partition function as:

$$\mathcal{Z} = \int e^{-\beta E + S(E)}\, dE \tag{12}$$

Is this relationship true for all $N$, or only in the thermodynamic limit $N \to \infty$?

We have:

$$\mathcal{Z} = \int e^{-\beta\mathcal{H}[s]}\, ds = \int e^{-\beta E}\delta(\mathcal{H}[s] - E)\, dE ds = \int e^{-\beta E}\, dE \int \delta(\mathcal{H}[s] - E)\, ds$$

The result follows from the definition of $S(E)$:

$$S(E) = \log\left[\int \delta(\mathcal{H}[s] - E)\, ds\right]$$

3. Combine the last two results to argue that in the large $N$ limit:

$$S(E_{\mathrm{eq}}) = \Phi(E_{\mathrm{eq}}) + \beta E_{\mathrm{eq}} \tag{13}$$

where $E_{\mathrm{eq}}$ is the energy given from the saddle point approximation maximisation condition. What is the condition that determines $E_{\mathrm{eq}}$?

Notice that both $E$ and $S(E)$ scale linearly with $N$, so in the large $N$ limit $e$ and $s(e)$ are finite (and can take non-zero values):

$$e = \frac{E}{N} \qquad s(e) = \frac{S(e)}{N}$$

We can now apply the saddle point method on equation (12) to obtain that up to a constant independent from $N$ we have:

$$\Phi = -\beta E_{\mathrm{eq}} + S(E_{\mathrm{eq}})$$

where

$$E_{\mathrm{eq}} = \arg\min_{E}(S(E) - \beta E)\,, \tag{14}$$

which gives the condition

$$\beta = \frac{\partial S(E)}{\partial E}\,. \tag{15}$$

Similarly in this limit $\langle E\rangle = E_{\mathrm{eq}}$ and $\Phi = \Phi(E_{\mathrm{eq}})$.

## Central Limit Theorem using field theory

Consider $N$ independent samples $x_1$, ... $x_N$ of a random variable $x \sim p(x)$, where $p$ has mean $\mu$ and variance $\sigma^2$. We define the random variable $y_N$ as the average of the $N$ samples:

$$y_N = \frac{1}{N}\sum_n x_n\,, \tag{16}$$

then in the large $N$ limit $y_N$ converges in distribution to $y \sim q(y) = \mathcal{N}(\mu, \sigma^2/N)$. This basic fact of probability theory is called the Central Limit Theorem. As an exercise we will derive it using field theory techniques.

Recall the definition of the Dirac's delta distribution $\delta$

$$\int dx \delta(x - x_0) f(x) = f(x_0)\,, \tag{17}$$

and its Fourier representation

$$\delta(x) = \int \frac{d\hat{x}}{2\pi} \exp^{i\hat{x}x}\,. \tag{18}$$

1. Write the distribution of $y_n$, which we want to show it converges to $q(y)$, as a function of $p(x)$ by using the delta function to impose the definition of $y$.

$$q(y) = \int \delta\left(\sum_n x_n/N - y\right) \prod_n p(x_n)\,\mathrm{d}x_n \tag{19}$$

2. Rewrite the delta in Fourier representation (also called informally exponential form).

$$q(y) = \int \mathrm{d}\hat{y}\, e^{-\hat{y}y + \hat{y}\sum_n x_n/N} \prod_n p(x_n)\,\mathrm{d}x_n \tag{20}$$

   where $\hat{y}$ is integrated along the imaginary line.

3. We start with a weaker form of the result (the law of large numbers): let's show that at the zeroth order in $N$, $\hat{y}_N$ converges in distribution to $q(y) = \delta(y - \mu)$. Do it by expanding the exponential in power series, keep the zeroth order terms in $N$, then resum.

   We will only write the relevant piece

$$\int e^{\hat{y}\sum_n x_n/N} \prod_n p(x_n)\,\mathrm{d}x_n \approx \tag{21}$$

$$\int \left[\sum_{k=0}^{\infty} \frac{\hat{y}^k}{k!N^k} \left(\sum_n x_n\right)^k\right] \prod_n p(x_n)\,\mathrm{d}x_n \approx \tag{22}$$

$$\int \left[\sum_{k=0}^{\infty} \frac{\hat{y}^k}{k!N^k} \sum_{n_1 \neq \ldots \neq n_k} x_{n_1} \ldots x_{n_k}\right] \prod_n p(x_n)\,\mathrm{d}x_n \approx \tag{23}$$

$$\int \sum_{k=0}^{\infty} \frac{\hat{y}^k}{k!} \mu^k = \tag{24}$$

$$e^{\hat{y}\mu} \tag{25}$$

   Here we suppressed all the subleading pieces in the sum: we need $N^k$ elements to be summed, so all the indices need to be different. Then, we can integrate all terms $x_{n_i}$ independently obtaining the mean $\mu$, and argue that

$$N^{-k} \sum_{n_1 \neq \ldots \neq n_k} 1 = 1 \tag{26}$$

   at leading order. We can now plug this in our expression for $q(y)$.

$$q(y) = \int \mathrm{d}\hat{y}\, e^{\hat{y}(\mu - y)} = \delta(y - \mu) \tag{27}$$

4. (Bonus) As we saw from the previous computation, $y = \mu$ at leading order in $N$. Thus, in the previous computation, we could have avoided enforcing the definition of $y$ using the $\delta$ distribution, as $y$ naturally respects the constraint enforced by the delta in the large $N$ limit (recall that $\sum_i x_i/N \to \mu$ for large $N$). Whenever this is the case, i.e. whenever we enforce a "vacuous" constraint using a delta function, we can take $\hat{y} \approx 0$. Intuitively, $\hat{y}$ is an external field that enforces the constraint (very much like a magnetic field used to induce a magnetisation in a magnetic system), and if the system satisfies already the constraint, no external field is needed.

   Thus, it's reasonable to expand around $\hat{y} = 0$. Expand the exponential in power series and keep only the leading order terms up to second order in $\hat{y}$, then "resum" the exponential to show that the fluctuations are Gaussian

   We will only write the relevant piece

$$\int e^{\hat{y}\sum_n x_n/N} \prod_n p(x_n)\, \mathrm{d}x_n \approx \tag{28}$$

$$\int \left[1 + \frac{\hat{y}}{N}\sum_n x_n + \frac{\hat{y}^2}{2N^2}\left(\sum_n x_n\right)^2\right] \prod_n p(x_n)\, \mathrm{d}x_n = \tag{29}$$

$$\int \left[1 + \frac{\hat{y}}{N}\sum_n x_n + \frac{\hat{y}^2}{2N^2}\left(\sum_n x_n^2 + \sum_{n\neq m} x_n x_m\right)\right] \prod_n p(x_n)\, \mathrm{d}x_n = \tag{30}$$

$$1 + \hat{y}\mu + \frac{\hat{y}^2\sigma^2}{2N} + \frac{\hat{y}^2\mu^2}{2} \tag{31}$$

   This is again the same at leading order to the exponential

$$e^{\hat{y}\mu + \frac{\hat{y}^2\sigma^2}{2N}} \tag{32}$$

   We are left with

$$\int \mathrm{d}\hat{y}\, e^{\hat{y}(y-\mu) + \frac{\hat{y}^2\sigma^2}{2N}} \tag{33}$$

   Which is exactly a Gaussian distribution after integrating over $\hat{y}$.