Mathematical Statistical Physics – LMU München, summer semester 2012

Hartmut Ruhl, Imre Péter Tóth

Homework sheet 9 – due on 22.06.2012 – and exercises for the class on 15.06.2012

9.1 **(homework)** *Canonical partition function and density of states.* As you all know, the canonical ensamble (or canonical distribution) has the density $f_{\alpha,\beta}(\omega) = \frac{1}{Z(\alpha,\beta)} e^{-\beta H(\omega)}$ w.r.t. an approriate reference measure μ_{ref} on the phase space $\Omega = \{\omega\}$. Here β is the inverse temperature, and α denotes all the possibles other parameters (e.g. volume, particle number, etc.) which influence the shape of Ω, μ_{ref} and the Hamiltonian $H : \Omega \to \mathbb{R}$. The normalizing factor $Z(\alpha, \beta)$ is called the partition function (we suppose that it is finite).

Denote by μ_E the push-forward of μ_{ref} from Ω to $\mathbb{R}$ by H – which means that

$$\mu_E(B) := \mu_{ref}(\{\omega : H(\omega) \in B\})$$

for any Borel $B \subset \mathbb{R}$. This could vaguely be called the "distribution of H w.r.t. μ_{ref}". (Only vaguely, because μ_{ref} is usually not a probability measure, so $H : \Omega \to \mathbb{R}$ cannot be called a random variable if we consider Ω equipped with μ_{ref}.) Suppose (for simplicity only) that this μ_E is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}$, and denote its density by $\rho = \rho_\alpha(E)$. This ρ_α can be called the **density of states**.

(a) When Ω is equipped with the canonical measure, the energy is a random variable. Show that under the above condition (that μ_E is absolutely continuous w.r.t. Lebesgue measure) this random variable is absolutely continuous (w.r.t Lebesgue measure), and calculate the density in terms of ρ, Z and β.

(b) Express $Z(\alpha, \beta)$ with the help of β and ρ_α (or β and μ_E, if you want to be more general), and be happy that this is possible.

Solution:

(a) Let $\mu_{can}^{\alpha,\beta}$ denote the canonical measure and let $\nu_{\alpha,\beta}^E$ denote the distribution of H w.r.t. $\mu_{can}^{\alpha,\beta}$, which is thus the push-forward of $\mu_{can}^{\alpha,\beta}$ by H (from ω to $\mathbb{R}$). So we can calculte it using the definition of the push-forward, the definition of the canonical measure and the theorem of integration by substitution: for any Borel $B \subset \mathbb{R}$

$$\begin{aligned} \nu_{\alpha,\beta}^E(B) &= \mu_{can}^{\alpha,\beta}(\{\omega : H(\omega) \in B\}) = \int_{H^{-1}(B)} \mathrm{d}\mu_{can}^{\alpha,\beta} = \\ &= \int_{H^{-1}(B)} \frac{1}{Z(\alpha,\beta)} e^{-\beta H(\omega)} \,\mathrm{d}\mu_{ref}(\omega) \overset{E=H(\omega)}{=} \int_B \frac{1}{Z(\alpha,\beta)} e^{-\beta E} \,\mathrm{d}\mu_E(E), \end{aligned}$$

so the general expression for $\nu_{\alpha,\beta}^E$ is

$$\mathrm{d}\nu_{\alpha,\beta}^E(E) = \frac{1}{Z(\alpha,\beta)} e^{-\beta E} \,\mathrm{d}\mu_E(E).$$

So if μ_E is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}$ with density ρ_α, then $\nu_{\alpha,\beta}^E$ also has a density

$$g_{\alpha,\beta}(E) = \frac{1}{Z(\alpha,\beta)} e^{-\beta E} \rho_\alpha(E).$$

(b) From the previous, the normalizing factor has to be, in the general case,

$$Z(\alpha, \beta) = \int_{\mathbb{R}} e^{\beta E} \, \mathrm{d}\mu_E(E).$$

In the special case when μ_E is absolutely continuous,

$$Z(\alpha, \beta) = \int_{-\infty}^{\infty} e^{\beta E} \rho_\alpha(E) \, \mathrm{d}E.$$

9.2 **(homework)** *Energy fluctuations for the free gas.* Consider the free gas in the canonical ensamble, and keep the density fixed by setting $V = Nv$ with $v = const.$ Also fix the temperature by setting $\beta = const.$ Now for every N the energy density H/V is a random variable.

(a) Calculate the expectation and the variance of this H/V as a function of N. What can we say about the weak convergence of H/V in the limit $N \to \infty$?

(b) Set $N = 10^{23}$. Estimate the probability that H/V deviates from its expectation with at least 0.000001%.

Solution1: brute force calculation, without understanding what the partition function is good for – but understanding what the canonical distribution is.

(a) $H = \frac{1}{2m} \sum_{i=1}^{3N} p_i^2$, where each p_i is one of the $3N$ moment vector components. In the canonical ensamble, these p_i are *random variables*, whose distribution is known exactly: they are i.i.d. and all of them are Gaussian with mean 0 and variance $\frac{m}{\beta}$. This information is enough to calcualte the expectation and variance of H: linearity of the expectation implies that

$$\mathbb{E}H = \frac{1}{2m} 3N \mathbb{E}(p_1^2),$$

and inpedendence implies that

$$\mathrm{Var}H = \frac{1}{(2m)^2} 3N \mathrm{Var}(p_1^2).$$

$\mathbb{E}(p_1^2)$ and $\mathrm{Var}(p_1^2)$ can be calculated using only the fact that $p_1 \sim \mathcal{N}(0, \frac{m}{\beta})$:

$$\mathbb{E}(p_1^2) = \mathrm{Var} p_i = \frac{m}{\beta}$$

and

$$\mathbb{E}((p_1^2)^2) = \mathbb{E}(p_1^4) = \int_{-\infty}^{\infty} x^4 \frac{1}{\sqrt{2\pi \frac{m}{\beta}}} e^{-\frac{x^2}{2\frac{m}{\beta}}} \, \mathrm{d}x = \cdots = \frac{3m^2}{\beta^2},$$

so

$$\mathrm{Var}(p_1^2) = \mathbb{E}((p_1^2)^2) - (\mathbb{E}(p_1^2))^2 = \frac{2m^2}{\beta^2}.$$

So

$$\mathbb{E}H = \frac{3N}{2\beta} \quad , \quad \mathrm{Var}H = \frac{3N}{2\beta^2}.$$

Now using $V = Nv$ we get

$$\mathbb{E}\frac{H}{V} = \frac{3}{2v\beta} \quad , \quad \mathrm{Var}\frac{H}{V} = \frac{3}{2v^2\beta^2} \frac{1}{N}.$$

So, as a function of N, the expectation is constant and the variance goes to zero, which ensures that $\frac{H}{V}$ converges to $\frac{3}{2v\beta}$ weakly as $N \to \infty$.

(b) i. *Easiest, **very** rough estimate using the Markov (or the Chebyshev's) inequality:* Use the notation $\delta = 10^{-8}$.

$$\begin{aligned}\mathbb{P}\left(\left|\frac{H}{V} - \mathbb{E}\left(\frac{H}{V}\right)\right| > \delta\mathbb{E}\left(\frac{H}{V}\right)\right) &= \mathbb{P}\left(\left(\frac{H}{V} - \mathbb{E}\left(\frac{H}{V}\right)\right)^2 > \delta^2\mathbb{E}^2\left(\frac{H}{V}\right)\right) \leq \\ &\leq \frac{\mathrm{Var}\left(\frac{H}{V}\right)}{\delta^2\mathbb{E}^2\left(\frac{H}{V}\right)} = \frac{2\delta^2}{3N} = 6.666 \cdot 10^{-8}.\end{aligned}$$

ii. *Much better estimate using large deviations:* If we write H in the form $H = \sum_{i=1}^{3N} X_i$ where $X_i = \frac{1}{2m}p_i^2$, we can give a large deviations estimate for

$$\mathbb{P}\left(\left|\frac{H}{V} - \mathbb{E}\left(\frac{H}{V}\right)\right| > \delta\mathbb{E}\left(\frac{H}{V}\right)\right) = \mathbb{P}\left(\left|\frac{H}{3N} - \mathbb{E}X\right| > \delta\mathbb{E}X\right)$$

by calcualting the Cramer rate function for $X := X_1$. For that, it's enough to know the distribution of $p = p_1$ and the definition of X: the moment generating function is

$$Z(\lambda) = \mathbb{E}(e^{\lambda X}) = \mathbb{E}(e^{\frac{\lambda}{2m}p^2}) = \int_{-\infty}^{\infty} e^{\frac{\lambda}{2m}x^2} \frac{1}{\sqrt{2\pi\frac{m}{\beta}}} e^{-\frac{x^2}{2\frac{m}{\beta}}}\,\mathrm{d}x = \cdots = \sqrt{\frac{\beta}{\beta-\lambda}}.$$

From that we get

$$\hat{I}(\lambda) = \log Z(\lambda) = \frac{1}{2}\log\beta - \frac{1}{2}\log(\beta-\lambda),$$

so $\mathbb{E}X = \hat{I}'(0) = \frac{1}{2\beta}$, $x = \hat{I}'(\lambda^*)$ gives $\lambda^*(x) = \beta - \frac{1}{2x}$, so

$$I(x) = x\lambda^*(x) - \hat{I}(\lambda^*) = x\beta - \frac{1}{2} - \frac{1}{2}\log(2\beta x)$$

and the Cramer theorem gives

$$\begin{aligned}\mathbb{P}\left(\frac{H}{3N} < (1-\delta)\frac{1}{2\beta}\right) &\lessapprox e^{-3NI(\frac{1-\delta}{2\beta})}, \\ \mathbb{P}\left(\frac{H}{3N} > (1+\delta)\frac{1}{2\beta}\right) &\lessapprox e^{-3NI(\frac{1+\delta}{2\beta})}.\end{aligned}$$

The essential part is

$$\begin{aligned}I(\frac{1-\delta}{2\beta}) &= \frac{1}{2}(-\delta - \log(1-\delta)) \approx \frac{\delta^2}{4}, \\ I(\frac{1+\delta}{2\beta}) &= \frac{1}{2}(\delta - \log(1+\delta)) \approx \frac{\delta^2}{4},\end{aligned}$$

and

$$\mathbb{P}\left(\left|\frac{H}{3N} - \mathbb{E}X\right| > \delta\mathbb{E}X\right) \lessapprox 2e^{-\frac{3N\delta^2}{4}} = 2e^{-7.5\cdot 10^6},$$

which has roughly 3257000 zeroes before the first significant digit.

Solution2: Short and easy calculation, making use of the partition function.

(a) In Exercise 5.4 we calculated the canonical partition function

$$Z(N,V,\beta) = \frac{V^N}{N!}\left(\frac{2\pi m}{\beta}\right)^{\frac{3N}{2}}, \text{ so } \log Z(N,V,\beta) = const(N,V) - \frac{3N}{2}\log\beta.$$

This implies

$$\mathbb{E}H = -\frac{\partial}{\partial\beta}\log Z(N,V,\beta) = \frac{3N}{2\beta} \text{ and } \mathrm{Var}H = \frac{\partial^2}{\partial\beta^2}\log Z(N,V,\beta) = \frac{3N}{2\beta^2}.$$

The rest is the same as in the first solution.

(b) i. *Easiest,* ***very*** *rough estimate using the Markov (or Chebyshev's) inequality:* same as in the first solution.

ii. *Much better estimate using large deviations:* The great thing in the definition of the partition function is exactly that $Z(N,V,\beta)$, as a function of β, is essentially the moment generating function of the random variable H. To be precise,

$$\begin{aligned}\mathbb{E}_{\mu_{can}}(e^{\lambda H}) &= \int_\Omega e^{\lambda H(\omega)}\,\mathrm{d}\mu_{can}(\omega) = \int_\Omega e^{\lambda H(\omega)}\frac{1}{Z(N,V,\beta)}e^{-\beta H(\omega)}\,\mathrm{d}\mu_{ref}(\omega) = \\ &= \frac{1}{Z(N,V,\beta)}\int_\Omega e^{-(\beta-\lambda)H(\omega)}\,\mathrm{d}\mu_{ref}(\omega) = \frac{1}{Z(N,V,\beta)}Z(N,V,\beta-\lambda).\end{aligned}$$

So, having already calculated the partition function, we get the moment generating function for free:

$$\mathbb{E}(e^{\lambda H}) = \left(\frac{\beta}{\beta-\lambda}\right)^{\frac{3N}{2}}.$$

To avoid confusion, let's denote the logarithmic moment generating function with $\hat{J}$:

$$\hat{J}(\lambda) := \log\mathbb{E}(e^{\lambda H}) = \frac{3N}{2}(\log\beta - \log(\beta-\lambda)).$$

(Note that this $\hat{J}$ is *not the same* as the $\hat{I}$ in the first solution: $\hat{I}$ denoted the logarithmic moment generating function of X, while $\hat{J}$ is the logarithmic moment generating function of H. Of course, $\hat{J}(\lambda) = 3N\hat{I}(\lambda)$.)

We will simply estimate $\mathbb{P}(|H-\mathbb{E}H| \geq \delta\mathbb{E}H)$ using the large deviations theorem with $n=1$ – that is, for a sum with the single term H. For the rate function we get

$$J(x) = x\beta - \frac{3N}{2} - \frac{3N}{2}\log\frac{2\beta x}{3N}.$$

(Note that this is related naturally to the rate function of the previous solution: $J(x) = 3NI(\frac{x}{3N})$.)

The Cramer theorem gives

$$\begin{aligned}\mathbb{P}(H < (1-\delta)\mathbb{E}H) &\lessapprox e^{-J\left((1-\delta)\frac{3N}{2\beta}\right)} = e^{-\frac{3N}{2}(-\delta-\log(1-\delta))} \approx e^{-\frac{3N\delta^2}{4}},\\ \mathbb{P}(H > (1+\delta)\mathbb{E}H) &\lessapprox e^{-J\left((1+\delta)\frac{3N}{2\beta}\right)} = e^{-\frac{3N}{2}(\delta-\log(1+\delta))} \approx e^{-\frac{3N\delta^2}{4}},\end{aligned}$$

so

$$\mathbb{P}(|H-\mathbb{E}H| \geq \delta\mathbb{E}H) \lessapprox 2e^{-\frac{3N\delta^2}{4}} = 2e^{-7.5\cdot 10^6},$$

small.

9.3 *Density fluctuations for the free gas.* Consider the free gas in the grand canonical ensamble. Keeping β and β' fixed, the density N/V is a random variable parametrized by V.

(a) Calculate the expectation and the variance of this N/V as a function of V. What can we say about the weak convergence of N/V in the limit $V \to \infty$?

(b) Set the parameters so that $\mathbb{E}N = 10^{23}$. Estimate the probability that N/V deviates from its expectation with at least 0.000001%.

9.4 *Tempered and stable pair interactions.* Let $\Phi : \mathbb{R}^+ \to \mathbb{R} \cup \{\infty\}$ be a pair interaction potential which satisfies the following:

(a) Φ is bounded from below,

(b) There is an $R_1 > 0$ such that $\Phi(r) = \infty$ for all $r \leq R_1$,

(c) There is an $R_2 < \infty$ such that $\Phi(r) = 0$ for all $r \geq R_2$.

Show that Φ is tempered and stable.

9.5 *Tempered and stable pair interactions II.* Let $\Phi : \mathbb{R}^+ \to \mathbb{R} \cup \{\infty\}$ be a pair interaction potential which satisfies the following:

(a) Φ is bounded from below,

(b) There is an $R_1 > 0$ such that $\Phi(r) = \infty$ for all $r \leq R_1$,

(c) There is an $R_2 < \infty$ such that $\Phi(r) \leq 0$ for all $r \geq R_2$,

(d) $\Phi(r) \to 0$ exponentially fast as $r \to \infty$.

Show that Φ is tempered and stable.

9.6 *Basics of convex functions.* If a and b are elements of a linear space V over $\mathbb{R}$, then their **convex combinations** are the elements $\alpha a + \beta b$ where $0 \leq \alpha \in \mathbb{R}$, $0 \leq \beta \in \mathbb{R}$ and $\alpha + \beta = 1$. A **set** $A \subset V$ is called **convex** if it contains every convex combination of its elements. For a convex $A \subset V$, the **function** $f : A \to \mathbb{R} \cup \{\infty\}$ is called **convex** if

$$f(\alpha a + \beta b) \leq \alpha f(a) + \beta f(b)$$

for any $a, b \in A$, $0 \leq \alpha \in \mathbb{R}$, $0 \leq \beta \in \mathbb{R}$ and $\alpha + \beta = 1$. Show that convexity is a very strong regularity property by proving the following statements: Suppose $f : I \to \mathbb{R} \cup \{\infty\}$ is convex and finite on the open (but possibly infinite) interval $I \subset \mathbb{R}$. Then

(a) it is necessarily continuous,

(b) it has one-sided derivatives everywhere on I,

(c) These one-sided derivatives are monotonically non-decreasing,

(d) f is differentiable in all but at most countably many points.

9.7 *Midpoint convexity.* Let $I \subset \mathbb{R}$ be a (possibly infinite) interval. The function $f : I \to \mathbb{R} \cup \{\infty\}$ is called **midpoint convex**, if $f(\frac{a+b}{2}) \leq \frac{f(a)+f(b)}{2}$ for every $a, b \in I$. Show that if $f : I \to \mathbb{R} \cup \{\infty\}$ is finite, midpoint convex and bounded on a subinterval $\emptyset \neq J \subset I$, then it is bounded on any bounded interval, (continuous) and convex.

9.8 **(homework)** *Jensen's inequality.* If $a_1, \ldots, a_n$ are elements of a linear space V over $\mathbb{R}$, then their convex combinations are the elements $\sum_{i=1}^n \alpha_i a_i$ where $0 \leq \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $\sum_{i=1}^n \alpha_i = 1$.

(a) Show that if $A \subset V$ is convex and $a_1, \ldots, a_n \in A$, then any convex combination $\sum_{i=1}^{n} \alpha_i a_i$ is also in A.

(b) Show that if $A \subset V$ is convex, $f : A \to \mathbb{R} \cup \{\infty\}$ is convex and $a_1, \ldots, a_n \in A$, $0 \leq \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $\sum_{i=1}^{n} \alpha_i = 1$, then

$$f\left(\sum_{i=1}^{n} \alpha_i a_i\right) \leq \sum_{i=1}^{n} \alpha_i f(a_i).$$

This is the simplest form of Jensen's inequality.

Solution: Very easy, by induction in n. For $n = 1$ the statements are trivial identities, for $n = 2$ they are the definitions of convexity (of A and of f, respectively). For $n \geq 3$ assume that the statements hold for $n - 1$.

Set $\beta_1 = \sum_{i=1}^{n-1} a_i$ and $\beta_2 = a_n$, so $\beta_1, \beta_2 \geq 0$ and $\beta_1 + \beta_2 = 1$.

If $\beta_1 = 0$, the statements are trivial. If not, set $\gamma_i = \frac{\alpha_i}{\beta_1}$ for $i = 1, \ldots, n-1$, so $\sum_{i=1}^{n-1} \gamma_i = 1$. Set $P := \sum_{i=1}^{n} a_i$ and $b_1 := \sum_{i=1}^{n-1} \gamma_i a_i$.

Now

(a) The inductive assumption implies that $b_1 \in A$, so the convexity of A implies that $\beta_1 b_1 + \beta_2 a_n \in A$, but $\beta_1 b_1 + \beta_2 a_n = P$. $\square$

(b) The convexity if f implies that $f(P) = f(\beta_1 b_1 + \beta_2 a_n) \leq \beta_1 f(b_1) + \beta_2 f(a_n)$, and the inductive assumption implies that $f(b_1) = f(\sum_{i=1}^{n-1} \gamma_i a_i) \leq \sum_{i=1}^{n-1} \gamma_i f(a_i)$. Putting these together,

$$f(P) \leq \beta_1 \sum_{i=1}^{n-1} \gamma_i f(a_i) + \beta_2 f(a_n) = \sum_{i=1}^{n} \alpha_i f(a_i).$$

$\square$