Statistical Physics IV: Non-equilibrium statistical physics

ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE (EPFL)

Exercise No.7

7.1 Master equation for an asymmetric random walk

Consider a random walk discrete in space and continuous in time. The probability rates to jump from n to $n \pm 1$ are given by α and β, respectively.

1. Write the master equation. Derive an equation of motion for the generating function $G(z,t) = \sum_n z^n P(n,t)$.
2. Find a solution for the generating function with the initial condition $P(n,0) = \delta_{n_1,n}$ (the system is initially in the state n_1 at time $t = 0$).
3. Show that for the fully asymmetric case $\beta = 0$, we get the Poisson distribution $P(n_2,t) = e^{-\alpha t}\frac{(\alpha t)^{n_2-n_1}}{(n_2-n_1)!}$ for $n_2 \geq n_1$ and 0 for $n_2 < n_1$.

7.2 Equilibrium between photons and atoms

Consider an ensemble of atoms with two states E_1 and E_2, resonant with a mode of the radiation field of frequency $\omega = |E_2 - E_1|/\hbar$. The numbers of atoms in the ground and excited state are N_1 and N_2. According to the quantum theory of light the energy of the radiation field is quantized and obeys $E = \hbar\omega n$, where n is the number of photons in the radiation field. Transitions from the ground state to the excited state of the atoms absorb a photon, and happen at a rate $r_n = n\gamma N_1$. Transition from the excited to the ground state create a photon, and happen at a rate $g_n = (n+1)\gamma N_2$. The extra factor of one originates from *spontaneous emission*. We assume the numbers of atoms in each state N_1 and N_2 to be fixed by some other process and stay constant, so that we only consider the fluctuations of the photon number n.

1. Derive the master equation and solve it in the steady state.
2. Assume that the atoms are Ammonia for instance (as used in a maser[1]), whose energy spacing is in the GHz range (thus at room temperature $k_B T \gg |E_2 - E_1| = \hbar\omega$). If the atoms are in thermal equilibrium, it follows from the Boltzmann statistics that $N_2/N_1 = e^{-\hbar\omega/k_B T}$. Derive an expression for the steady state distribution of P_n^s and calculate $\langle n \rangle$. Show that this yields the Bose-Einstein statistics.

7.3 Master equation for a chemical reaction (Dimer formation)

Assume you have a reservoir with influx of monomers at a rate K_+. The monomers that enter the reservoir can also exist with a rate K_-(with unit inverse monomers per time). Assume that in the reservoir the monomers can react to form dimers at a rate $2K_D$ (with unit inverse monomers squared per time).

1. Show that the master equation for the number of monomers (n) is given by:

$$\frac{\partial p_n(t)}{\partial t} = K_+ p_{n-1}(t) + K_-(n+1)p_{n+1}(t) + K_D(n+2)(n+1)p_{n+2}(t) - [K_+ + K_- n + K_D n(n-1)]p_n(t) \tag{1}$$

   and explain what each term in the right hand side of the equation corresponds to.

[1] **M**icrowave **a**mplification by **s**timulated **e**mission of **r**adiation. The maser was was the first oscillator providing a coherent source of electromagnetic radiation. Later worked showed the principle could be extended to the visible domain, leading to the Laser.

2. By taking the moments of the Master equation show that the mean number of molecules that have not formed molecules (mono-mers) evolve according to the deterministic rate equation:

$$\frac{d}{dt}\langle n\rangle = K_+ - K_-\langle n\rangle - 2K_D\langle n(n-1)\rangle.$$

3. Show that the generating function $G(z,t) = \sum_0^\infty z^n P_n(t)$ satisfies the equation:

$$\frac{\partial}{\partial t}G(z,t) = K_D(1-z^2)\frac{\partial^2}{\partial z^2}G(z,t) + K_-(1-z)\frac{\partial}{\partial z}G(z,t) - K_+(1-z)G(z,t) \tag{2}$$

4. Given the steady solution of this differential equation, $G(z) = (\frac{z+1}{2})^{(1-2\beta)/2} \times \frac{I_{2\beta-1}[\sqrt{8\alpha(z+1)}]}{I_{2\beta-1}[\sqrt{16\alpha}]}$, in the limit of small fluxes, numerically find the value of mean $\langle n\rangle$ and standard deviation $\sigma = \sqrt{\langle n^2\rangle - \langle n\rangle^2}$ of the monomer number. Assume values of $\alpha = 0.1, \beta = 1$ as well as $\alpha = 1, \beta = 0.1$. Here $\beta = \frac{K_-}{2K_D}$, $\alpha = \frac{K_+}{2K_D}$ and I_v denotes the modified Bessel function of the first kind of order v.

5. For the rest of the exercise we consider the reaction in steady state. Qualitatively, under what conditions (expressed as relations between K_+, K_- and K_D) do the relative fluctuations in the monomer number n become small? (hint: fluctuations are small in macroscopic limit, when average particle numbers are large)

6. Assuming that the dimer formation rate $K_D\langle n(n-1)\rangle$ is small compare to both the influx K_+ and the loss $K_-\langle n\rangle$ of monomers, what should be the value of K_-/K_+ in order to keep the fluctuations $\sigma/\langle n\rangle$ in the monomer number below 10%?

7.4 The use of the Monte Carlo Metropolis in Bayesian statistical analysis: Spectral line problem[2]*

This problem illustrates the close connection and the use of Monte Carlo Metropolis algorithm, originally conceived to compute probability distributions of thermodynamical configurations[3] in the context of statistical data analysis, notably Baysian statistical analysis (the method referred to as Monte Carlo Markov Chains, MCMC). Baysian statistical analysis is a powerful method used in parameter estimation or model validation. At the heart of the method lies the fast sampling of the posteriori probability distribution, which is often computationally costsly, and therefore relies on numerical simulations (for which various software tools are widely available, see for example EMCEE[4]).

In this problem, we analyze the distinguishability of a spectral line from the Orion nebula detected with an astrophysical spectrometer. Given the spectrometer data, we have two models (M_1 and M_2), one in which a line exists with prior established bounds on its strength and a second one, in which no such line exists. Prior estimates of the line strength expected according to theory 1 range from $T_{min} = 0.1$mK to $T_{max} = 100$mK.

Theory 1 also predicts the line will have a Gaussian line shape of the form

$$T \cdot \exp\left\{\frac{-(\nu_i - \nu)^2}{2\sigma^2}\right\} = T \cdot f_i \tag{3}$$

where the signal strength is measured in temperature units of mK and T is the amplitude of the line. The frequency, ν_i, is in units of frequency channel number.

[2] cf. Chapters 3.6 and 12.6 of "Bayesian Logical Data Analysis for the Physical Sciences" P. C. Gregory
[3] Equation of State Calculations by Fast Computing Machines, N. Metropolis et. al. 1952
[4] GitHub repository: https://github.com/dfm/emcee

To test this prediction, a new spectrometer was mounted on the James Clerk Maxwell telescope on Mauna Kea and the spectrum shown in Figure 1 (Figure 3.3 from the referenced book) was obtained. The spectrometer has 64 frequency channels. All channels have Gaussian noise charac-

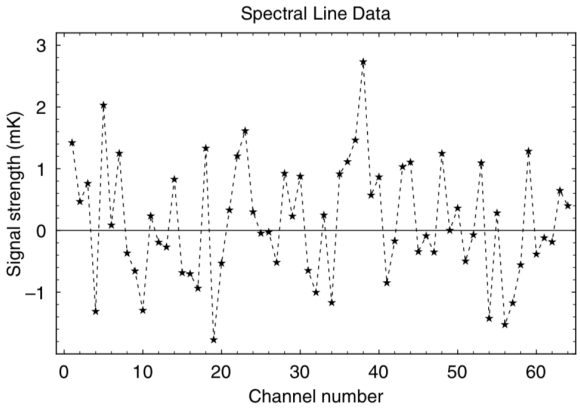

Figure 1: Measured spectrum

terised by $\sigma = 1$mK and the noise in separate channel is independent so that each acquired data is in the form

$$d_i = T \cdot f_i + e_i \tag{4}$$

where e_i is a zero mean Gaussian random variable with variance $\sigma = 1$mK.
This theory (or model) then is characterized by two parameters: $Y = \{T, \nu\}$, and we are therefore faced with a parameter estimation problem.The aim here is using Metropolis Hastings to estimate marginal posteriors of the line strength and center frequency, which are probability distributions of parameters T and ν_0, given that our theory, M_1, is true, some prior knowledge about the distributions before having any data set and finally a data set on measurement of the spectral lines strength (given on the table 3.1 of the book cited). The credible regions of the estimated parameters can then be computed by "marginalization" of the joint probability distribution of the parameters (e.g. the 95 percent credible regions).

According to Bayes theorem for posteriors we get

$$P(Y|D, M_1, I) = \frac{P(Y|M_1, I).P(D|M_1, Y, I)}{P(D|M_1, I)} \tag{5}$$

- D: Data set
- Y: $\{T, \nu\}$ (model parameters).
- $P(Y|M_1, I)$: Prior distributions (prior information).
- $P(D|M_1, Y, I)$: Global likelihood function (probability of the entire data set, given the prior information and theory M_1.
- $P(D|M_1, I)$: Global likelihood (probability of the entire data set, given the prior information and theory $M_1 = \int dY P(D|M_1, Y, I)$.

For the priors we choose uniform distributions for ν in the range channel 1 to 64. And for T we use the so-called Jeffreys prior (which is a equal propability per decade on a logarithmic scale) given by

$$P(T|M_1, I) = \frac{1}{T \ln(T_{max}/T_{min})} \tag{6}$$

and the joint prior $P(Y|M_1, I)$ is given by $P(\nu|M_1, I)P(T|M_1, I)$.
The global likelyhood can be directly computed given our model of the noise and is given by

$$P(D|M_1, Y, I) = \frac{1}{\sqrt{(2\pi\sigma^2)^N}} \exp\left\{-\frac{\sum_i (d_i - Tf_i)^2}{2\sigma^2}\right\} \tag{7}$$

where d_i is acquired data from channel i. We would like to find the optimum choice for T and ν which gives the maximum posteriori probability. With defining $\Lambda(Y) = -\ln(P(Y|M_1, I) \cdot P(D|M_1, Y, I))$, the posterior can be rewritten as

$$P(Y|D, M_1, I) = \frac{e^{-\Lambda(Y)}}{Z} \tag{8}$$

with $Z = \int dY e^{-\Lambda(Y)}$, which is equal to Global likelihood.

Next we need to maximize the likelihood over model parameters using a numerical method, we choose to use Monte-Carlo approach in this problem. Recasting given by Eq. 8 essentially suggests using Metropolis algorithm with the Metropolis rate

$$r = e^{-(\Lambda(Y') - \Lambda(Y))} \tag{9}$$

for transition from Y to Y'. Having all these information, The Metropolis algorithm steps can be given as follows:

- Initialize X_0; set $t = 0$. In this example we set $X_0 = \{T_0 = 5, \nu_0 = 30\}$
- Repeat {
  - Obtain new sample Y from $q(Y|X_t)$
    $Y = \{T', \nu'\}$
    we set $q(T'|T_t) = \mathcal{N}(T_t, \sigma_T = 1.0)$ (Normal Distribution)
    and $q(\nu'|\nu_t) = \mathcal{N}(\nu_t, \sigma_f = 1.0)$ (Normal Distribution)
  - Compute the Metropolis ratio:

$$r = \frac{P(Y|D, M_1, I)}{P(X_t|D, M_1, I)} = \frac{P(T', \nu'|M_1, I)P(D|M_1, T', \nu', I)}{P(T_t, \nu_t|M_1, I)P(D|M_1, T_t, \nu_t, I)} \tag{10}$$

    Note: if T', ν' lie outside the prior boundaries set $r = 0$.
  - if $r > 1$ accept $X_{t+1} = Y$
    if $r < 1$, choose U from a uniform random distribution between 0 and 1.$U \sim Unif(0, 1)$.
  - Accept $X_{t+1} = Y$ if $U \leq r$, otherwise set $X_{t+1} = X_t$
  - increment t.}

Note the transition rate, $W(X_{t+1}, X_t)$, defined as the probability of transfer from state X_t to state X_{t+1} during one time step, in our algorithm can be written as

$$W(X_{t+1}, X_t) = min(1, r) \tag{11}$$

Note that the transition rates satisfy detailed balance, i.e.

$$P(X_t|D, M_1, I)W(X_{t+1}, X_t) = P(X_{t+1}|D, M_1, I)W(X_t, X_{t+1}) \tag{12}$$

Therefore simulation of the trajectories in the parameter space will lead to a distribution that resembles the posteriori PDF.

1. Simulate the algorithm above using programming language of your choice and find the optimum values for T and ν. Plot the posterior distributions for T and ν by plotting a histogram of all the obtained values for T and ν.

2. Use Eq.(5) for calculating the posterior distributions again. Compare to results from last part.

3. An allowed range for a parameter X with probability content C (e.g. $C = 0.95$) is provided by the *credible region* R_X defined as

$$\int_{R_X} dX P(X|D, M) = C, \tag{13}$$

with the posterior density inside R_X everywhere greater than outside it. Calculate R_T and R_ν for $C = 0.95$.