Advanced Statistical Physics: Quick Guide

J.Pearson*

(Dated: January 28, 2009)

Abstract

This is a quick guide – a summary – of the Advanced Statistical Physics course at the University of Manchester, taught by A.McKane between Sept '08 and Dec '08. These summary notes are based upon his lecture notes. A copy of the full lecture notes, on this topic, may be found at `www.jpoffline.com`.

Keywords: Stochastic dynamics, Markov processes, Markov chains, Focker-Planck equation, Langevin equation

*Electronic address: `jon@jpoffline.com`

Contents

I. PROBABILITY THEORY

The probability of an event m occurring, given k has, is given by the **conditional probability**

$$P(m|k) = \frac{P(m)}{P(k)}.$$

From this, we derive **Bayes' theorem**,

$$P(A \cup B) = P(A|B)P(B).$$

A **stochastic random variable** is some function X that sends the sample space to the space of possible outcomes; for example:

$$X : S \longmapsto x \in \mathbb{R}.$$

A **probability density function** $P(x)$ is the probability that a measurement will result in value x. It has the following properties:

$$\sum_i P(x_i) = 1, \quad P(x_i) \geq 0, \quad \forall i.$$

Moments are defined as

$$\begin{aligned} \mu_n = \langle X^n \rangle &= \sum_i x_i^n P(x_i) \\ &= \int dx\, x^n P(x). \end{aligned}$$

The Fourier transform of the **characteristic function** e^{ikx} gives the **generating function**

$$\begin{aligned} G(k) &= \int dx\, e^{ikx} P(x) \\ &= \sum_i P(x_i) e^{ikx_i} \\ &= \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \mu_n. \end{aligned}$$

Hence, the moments can be computed from the generating function, via

$$\mu_n = \frac{1}{i^n} \frac{d^n G}{dk^n}\bigg|_{k=0}.$$

The probability density function is just the inverse Fourier transform of the generating function,

$$P(x) = \int \frac{dk}{2\pi}\, G(k) e^{-ikx}.$$

Cumulants are defined by

$$\kappa_n = \frac{1}{i^n}\frac{d^n}{dk^n}\left(\ln G\right)\Big|_{k=0}.$$

And the first few are

$$\kappa_1 = \langle\langle X\rangle\rangle = \mu_1, \quad \kappa_2 = \left\langle\left\langle X^2\right\rangle\right\rangle = \mu_2 - \mu_1^2.$$

A **marginal probability density function** is computed via summing over unwanted variables;

$$\begin{aligned} P(x) &= \int dy\, P(x,y) \\ &= \sum_y P(x,y). \end{aligned}$$

Conditional means are computed via

$$\langle X\rangle_{Y=y} = \int dx\, x\, P(x|y),$$

with an interesting result being

$$P(x|y) = \langle\delta(X-x)\rangle_{Y=y}.$$

A **Gaussian distribution** is the probability density function that it entirely characterised by its first two cumulants;

$$P(x) = \frac{1}{\sqrt{2\pi\kappa_2}} e^{-\frac{(x-\kappa_1)^2}{2\kappa_2}}, \quad \kappa_n = 0, \quad n \geq 3.$$

The **central limit theorem** states that the probability density function of the mean of a stochastic variable goes to Gaussian, over an infinite ensemble average. This is only true if the variable has finite moments and moments are independent of n.

The **binomial distribution** is such that the probability of choosing n objects, from N, is

$$\binom{N}{n} p^n q^{N-n}, \quad q = 1-p,$$

and the expectation value of such a distribution is Np.

II. MARKOV PROCESSES

Such a process has a probability density function that depends only on the **current state of the system**, and not on previous times. Hence,

$$P(x_{k+1}, t_{k+1}|x_k, t_k; \ldots; x_1, t_1) = P(x_{k+1}, t_{k+1}|x_k, t_k).$$

Using this **Markov assumption** gives two equations

$$\begin{aligned} P(x_2,t_2) &= \int dx_1\, P(x_2,t_2|x_1,t_1)P(x_1,t_1), \\ P(x_3,t_3|x_1,t_1) &= \int dx_2\, P(x_3,t_3|x_2,t_2)P(x_2,t_2|x_1,t_1). \end{aligned}$$

The second equation is known as the **Chapman-Kolmogorov equation**, whereby all intermediate states are summed over. These are sometimes called the **Markov equations**.

A. Markov Chains

A Markov process, with **discrete state-space and time** is called a **Markov chain**. The first Markov equation can be written

$$P_n(t+1) = \sum_m Q_{nm}(t)P_m(t),$$

where we have identified

$$P_n(t) = P(n,t), \quad Q_{nm}(t) = P(n,t+1|m,t).$$

The matrix Q_{mn} is called a **stochastic matrix**.

As the stochastic matrix is (in general) non-Hermitian, we have right- and left-eigenstates

$$Q\left|\psi^{(i)}\right\rangle = \lambda^{(i)}\left|\psi^{(i)}\right\rangle, \quad \left\langle\chi^{(i)}\right| Q = \lambda^{(i)}\left\langle\chi^{(i)}\right|.$$

The states are orthonormal and complete,

$$\left\langle\chi^{(i)}|\psi^{(j)}\right\rangle = \delta_{ij}, \quad \sum_i \left|\psi^{(i)}\right\rangle\left\langle\chi^{(i)}\right| = 1$$

The eigenvalues are

$$\left|\lambda^{(i)}\right| \leq 1.$$

The sum of the elements in a column of a stochastic matrix is unity;

$$\sum_n Q_{nm} = 1,$$

which corresponds to the statement that something must happen in a system.

The **stationary state** is the state which does not depend upon time. Hence, it corresponds to a unit eigenvalue

$$Q\left|\psi^{(\mathrm{st})}\right\rangle = \lambda^{(\mathrm{st})}\left|\psi^{(\mathrm{st})}\right\rangle = \left|\psi^{(\mathrm{st})}\right\rangle.$$

Hence, the stationary state is such that

$$\lambda^{(\mathrm{st})} = 1, \quad \left|\psi^{(\mathrm{st})}\right\rangle, \quad \left\langle\chi^{(\mathrm{st})}\right| = 1.$$

The solution to the Markov chain equation is given by

$$P(t) = Q^t P(0),$$

where powers of the stochastic matrix are computed via

$$Q^t = \sum_i \lambda^{(i)^t} \left|\psi^{(i)}\right\rangle \left\langle\chi^{(i)}\right|.$$

The **late time behaviour** of a system is described by the right-eigenstates corresponding to the eigenvalue whose modulus is closest to unity. If two eigenvalues are the same, the system flicks between the two states.

Upon constructing stochastic matrices, one must be very careful in imposing boundary conditions.

B. Master Equation

A Markov process, with **discrete state-space and continuous time** has its dynamics described by the **master equation**

$$\frac{dP_n}{dt} = \sum_{n' \neq n} w_{nn'}(t) P_{n'}(t) - \sum_{n' \neq n} w_{n'n}(t) P_n(t).$$

The coefficient $w_{n'n}(t)$ is the rate of probability for transition from state n to state n'. Hence, we read the above as describing that the rate of change of the probability of a particular state is the sum over all transitions in, minus transitions out of that state.

A **one-step process** is one for whom systems only transition to states next to each other. That is, $w_{n'n} = 0$ unless $n' = n \pm 1$. We denote the possible transition rates as

$$w_{n+1,n} = g_n \quad w_{n-1,n} = r_n.$$

Hence, g_n is the rate of moving "up states", and r_n the rate of moving "down states". Using this, the **one-step master equation** becomes

$$\frac{dP_n(t)}{dt} = r_{n+1} P_{n+1}(t) + g_{n-1} P_{n-1}(t) - (g_n + r_n) P_n(t).$$

A **linear one-step process** is one for whom

$$r_n = an + b, \quad g_n = cn + d.$$

Linear one-step processes may be solved by introducing the generating function

$$F(z,t) = \sum_n z^n P_n(t).$$

Doing this turns the master equation into a partial differential equation in $F(z,t)$, which can be solved for $F(z,t)$, then inverted to find $P_n(t)$. Notice that

$$\langle n(t) \rangle = \left. \frac{\partial F}{\partial z} \right|_{z=1}.$$

The **macroscopic equation** can be found by multiplying the one-step master equation by $\sum n$. Thus, the macroscopic equation is

$$\frac{d}{dt} \langle n(t) \rangle = \langle g_n \rangle - \langle r_n \rangle .$$

If the master equation contains non-linear terms, the best we can do is to use the long-time approximation

$$\left\langle n^k \right\rangle \longrightarrow \langle n \rangle^k .$$

The **stationary state** is the state that is independent of time. Hence,

$$\frac{dP_n^{\mathrm{st}}}{dt} = 0.$$

Therefore, the one-step master equation becomes

$$r_{n+1} P_{n+1}^{\mathrm{st}} - g_n P_n^{\mathrm{st}} = r_n P_n^{\mathrm{st}} - g_{n-1} P_{n-1}^{\mathrm{st}}.$$

From which we define a conserved **probability current**

$$J_{n+1} = g_n P_n^{\mathrm{st}} - r_{n+1} P_{n+1}^{\mathrm{st}}.$$

Now, if the system has **reflecting boundaries**, then $J = 0$ everywhere (as it is conserved). Hence, for reflecting boundaries,

$$g_n P_n^{\mathrm{st}} = r_{n+1} P_{n+1}^{\mathrm{st}}.$$

This allows computation of all P_n^{st}, via

$$P_0^{\mathrm{st}} = \frac{1}{1 + \sum_{n>0} \frac{g_{n-1} g_{n-2} \cdots g_0}{r_n r_{n-1} \ldots r_1}}.$$

We can define, from the **master equation**, some matrix

$$W_{mn}(t) = w_{mn}(t) - \delta_{mn} \sum_{n' \neq n} w_{n'n}(t),$$

so that the master equation becomes

$$\frac{dP_n(t)}{dt} = \sum_m W_{nm}(t) P_m(t),$$

which is very similar to the Markov chains. Notice that W is not symmetric (generally).

Detailed balance is the condition that

$$w_{nn'} P^{\mathrm{st}}_{n'} = w_{n'n} P^{\mathrm{st}}_n.$$

If detailed balance holds, we can form the symmetric matrix

$$V_{nm} = \sqrt{\frac{P^{\mathrm{st}}_m}{P^{\mathrm{st}}_n}} W_{nm},$$

and also

$$\tilde{P}_n(t) = \frac{P_n(t)}{\sqrt{P^{\mathrm{st}}_n}},$$

so that the master equation reads

$$\frac{d\tilde{P}_n(t)}{dt} = \sum_m V_{nm}(t) \tilde{P}_m(t).$$

Hence, as V is symmetric, we can solve this easily. The **eigenequation** is

$$V \left|\phi^{(i)}\right\rangle = \mu^{(i)} \left|\phi^{(i)}\right\rangle, \quad \mu^{(i)} \leq 0.$$

Notice that the **stationary state** $\left|P^{(\mathrm{st})}\right\rangle$ has $\mu^{(\mathrm{st})} = 0$. Also notice that as the stochastic matrix V is symmetric, the left- and right-eigenstates are identical. The states form a complete set,

$$\sum_i \left|\phi^{(i)}\right\rangle \left\langle\phi^{(i)}\right| = 1.$$

The **formal solution** is

$$\left|\tilde{P}(t)\right\rangle = e^{Vt} \left|\tilde{P}(0)\right\rangle.$$

Using this, and noting that

$$e^{At} \left|\varphi^{(i)}\right\rangle = e^{a^{(i)}t} \left|\varphi^{(i)}\right\rangle,$$

the solution can be written

$$\tilde{P}_n(t) = \sum_{i,m} e^{\mu^{(i)}t} \phi_n^{(i)} \phi_m^{(i)} \tilde{P}_m(0).$$

This solution is only valid **if detailed balance holds**. The eigenstates of V relate to the right- and left-eigenstates of W via

$$\psi_n^{(i)} = \sqrt{P_n^{\mathrm{st}}}\phi_n^{(i)}, \quad \chi_n^{(i)} = \frac{\phi_n^{(i)}}{\sqrt{P_n^{\mathrm{st}}}}$$

The **stationary state** has

$$\psi_n^{(\mathrm{st})} = P_n^{\mathrm{st}} \quad \chi_n^{(\mathrm{st})} = 1, \quad \mu^{(\mathrm{st})} = 0.$$

The **long term behaviour** is determined by the state corresponding to the eigenvalue closest to zero.

III. DRIFT & DIFFUSION: FOCKER-PLANCK EQUATION

Markov systems which have **continuous state-space and time** can be described by the **Kramers-Moyal expansion**

$$\frac{\partial P}{\partial t} = \sum_{\ell=1}^{\infty} \frac{(-1)^\ell}{\ell!} \frac{\partial^\ell}{\partial x^\ell} \left[D^{(\ell)}(x,t) P(x,t) \right].$$

This can be truncated to give the **Focker-Planck equation** (FPE)

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x} \left[A(x,t) P(x,t) \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[B(x,t) P(x,t) \right].$$

The **jump moments** are defined such that

$$M_\ell(x,t,\Delta t) = D^{(\ell)}(x,t)\Delta t + \mathcal{O}\left(\Delta t\right)^2 = \Big\langle \left[x(t+\Delta t) - x \right]^\ell \Big\rangle_{x(t)=x}.$$

For **one-step processes**, we can find that

$$D^{(\ell)}(x,t) = \left(g_n + (-1)^\ell r_n \right).$$

The general method for getting a FPE from a system, is to rescale time and lengths such that we can make both the time and space step-size go to zero, without killing off the entire equation.

A. Forms of the FPE

We can introduce the **probability current**

$$J(x,t) = A(x,t)P(x,t) - \frac{1}{2}\frac{\partial}{\partial x}\left[B(x,t)P(x,t)\right],$$

so that the FPE takes on the form of a **continuity equation**

$$\frac{\partial P}{\partial t} + \frac{\partial J}{\partial x} = 0.$$

If the system has **reflecting boundaries** at $x = a, b$, then

$$\frac{\partial}{\partial t}\int_a^b dx\, P(x,t) = 0,$$

which is the statement of **conserved norm**.

We can write the RHS of the FPE as an operator, so that the FPE becomes

$$\frac{\partial P}{\partial t} = LP,$$

where

$$L = -\frac{\partial}{\partial x}A(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}B(x,t).$$

The **adjoint operator** is

$$L^\dagger = A(x,t)\frac{\partial}{\partial x} + \frac{1}{2}B(x,t)\frac{\partial^2}{\partial x^2}.$$

Notice that the **stationary state** corresponds to where $LP^{\rm st} = 0$. Hence,

$$A(x)P^{\rm st}(x) - \frac{1}{2}\frac{d}{dx}\left[B(x)P^{\rm st}(x)\right] = J = \text{const},$$

after assuming that $A(x), B(x)$ only. **Reflecting boundaries** allow this to be solved, as $J = 0$,

$$P^{\rm st}(x) = \frac{\mathcal{N}}{B(x)}e^{2\int_a^x dx' \frac{A(x')}{B(x')}}.$$

We can turn the FPE into a **Schrodinger-like** equation, by assuming that $B = \text{const}$. If we let

$$\tilde{P}(x,t) = \frac{P(x,t)}{\sqrt{P^{\rm st}(x)}},$$

then the FPE for $\tilde{P}(x,t)$ becomes

$$\frac{\partial \tilde{P}}{\partial t} = H\tilde{P},$$

where

$$H = \frac{B}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{B}U(x), \quad U(x) = \frac{1}{2}A^2(x) + \frac{B}{2}\frac{\partial A}{\partial x}.$$

This new operator is Hermitian $H = H^\dagger$. Notice that

$$H = \left(P^{\rm st}\right)^{-1/2} L \left(P^{\rm st}\right)^{1/2}.$$

If we let

$$Q(x,t) = \frac{P(x,t)}{P^{\rm st}(x)},$$

then Q satisifes the **adjoint equation**

$$\frac{\partial Q}{\partial t} = L^\dagger Q.$$

We can write the **time independent FPE**,

$$LP^{(\mu)} = -\mu P^{(\mu)},$$

whereby solutions are formed by a sum over eigenstates,

$$P(x,t) = \sum_\mu c_\mu P^{(\mu)}(x) e^{-\mu t}.$$

As L is non-Hermitian, need right- and left-eigenstates. We thus have

$$\left\langle Q^{(\mu')} | P^{(\mu)} \right\rangle = \delta_{\mu\mu'}.$$

By multiplying the solution by

$$\int dx \; |x\rangle \langle x| = 1,$$

we are able to deduce the following **equivalent orthonormality statements**

$$\int dx \, \frac{P^{(\mu')}(x) P^{(\mu)}(x)}{P^{\rm st}(x)} = \delta_{\mu\mu'},$$

$$\int dx \, P^{\rm st}(x) Q^{(\mu')}(x) Q^{(\mu)}(x) = \delta_{\mu\mu'},$$

$$\int dx \, \tilde{P}^{(\mu')}(x) \tilde{P}^{(\mu)}(x) = \delta_{\mu\mu'}.$$

The first two have weight functions.

Upon multiplying the solution by

$$\int dx Q^{(\mu')}(x),$$

we are able to show that

$$c_\mu e^{-\mu t} = \int dx\, Q^{(\mu)}(x) P(x,t), \quad \sum_\mu Q^\mu(x') P^{(\mu)}(x) = \delta(x - x').$$

The second of these expressions is the statement of **completeness**.

The **stationary state** has

$$\mu^{(\mathrm{st})} = 0, \quad Q^{(\mathrm{st})} = 1.$$

IV. STOCHASTIC DIFFERENTIAL EQUATIONS

A **Langevin equation** is such that

$$m\dot{v}_i = -\alpha v_i - \nabla_i V + \eta_i(t),$$

where $\eta_i(t)$ is some stochastic random variable. The **overdamped limit** of this equation has the $\dot{v}_i = 0$.

For $\eta_i(t)$ to be **white noise**, we impose the conditions

$$\langle \eta_i(t) \rangle = 0, \quad \langle \eta_i(t) \eta_j(t') \rangle = 2D\delta_{ij}\delta(t - t').$$

The Langevin equation is easiest to solve using an **integrating factor** method (if we take $V = 0$); the integrating factor will be

$$e^{-\int dt\, \alpha},$$

so that the Langevin equation reads

$$\frac{d}{dt}\left[e^{\int dt'\, \alpha} v(t) \right] = e^{\int dt'\, \alpha} \eta(t).$$

Upon solving for $v(t)$, we can compute various quantities:

$$\langle v(t) \rangle, \quad \left\langle v^2(t) \right\rangle, \quad \langle v(t_1) v(t_2) \rangle;$$

where the final quantity is a **correlation function**. A good trick when computing the correlation function, is to note that

$$\int_0^{t_1} dt' \int_0^{t_2} dt'' \longrightarrow \int_0^{t_1} dt' \left[\int_0^{t_1} dt'' + \int_{t_1}^{t_2} dt'' \right], \quad t_2 > t_1.$$

For example, this allows simplification of integrals such as

$$\begin{aligned}\int_0^{t_1} dt' \int_0^{t_2} dt''\, \delta(t'-t'')f(t',t'') &= \int_0^{t_1} dt' \left[\int_0^{t_1} dt''\, \delta(t'-t'')f(t',t'') + \int_{t_1}^{t_2} dt''\delta(t'-t'')f(t',t'')\right] \\ &= \int_0^{t_1} dt' \int_0^{t_1} dt''\, \delta(t'-t'')f(t',t'').\end{aligned}$$

A Langevin equation describes a Markov process if $\eta(t)$ is white noise, and the Langevin equation is of 1^{st} order. If the Langevin equation is of the form,

$$m\ddot{x} + \alpha\dot{x} + V'(x) = \eta(t),$$

where $\eta(t)$ is white noise, then this is not a Markov process. This is because

$$\ddot{x}(t) = \frac{x(t+\tau) + x(t-\tau) - 2x(t)}{\tau},$$

which requires "knowledge" of the previous state of the system, not just the state that it is in now; in contrast to

$$\dot{x} = \frac{x(t+\tau) - x(t)}{\tau}.$$

To turn it into a Markov process, just redefine such that

$$m\dot{v} + \alpha v + V' = \eta(t), \quad \dot{x} = v.$$

Then, this is two first order stochastic differential equations.

Langevin equations can be put into the form of a FPE by integrating them, and reading off moments.