5.B. General Theory

The following notations will be used for the various probability densities for the stochastic (or random) variable Y:

$$P_1(y_1, t_1) \equiv P_1(1) \equiv P(1) \equiv \text{probability density for } Y = y_1 \text{ at } t = t_1. \tag{5.1}$$

$$P_2(y_1, t_1 \,;\, y_2, t_2) \equiv P_2(1, 2) \equiv \text{joint probability density for } Y = y_1 \text{ at } t = t_1 \text{ and } Y = y_2 \text{ at } t = t_2. \tag{5.2}$$

$$P_n(y_1, t_1 \,;\, \dots \,;\, y_n, t_n) \equiv P_n(1, \dots, n) \equiv \text{joint probability density for } Y = y_1 \text{ at } t = t_1, \dots, \text{ and } Y = y_n \text{ at } t = t_n. \tag{5.3}$$

Obviously, every P_n is symmetric in its arguments:

$$P_n(1, 2, \dots, n) = P_n(2, 1, \dots, n) = \dots \tag{5.3a}$$

Note that all probabilities are non-negative:

$$P_n \geq 0 \tag{5.4}$$

Since the sum of all possible probabilities of an event is 1, we have

$$\int d\,y_1\, P_1(1) = 1 \qquad \forall\ t_1 \tag{5.6}$$

and more generally,

$$\int d\,y_n\, P_n(1, \dots, n) = P_{n-1}(1, \dots, n-1) \qquad \forall\ t_n \tag{5.5}$$

The n^{th} moment (or correlation function) of Y is defined as

$$\langle\, y(t_1)\, y(t_2) \dots y(t_n) \,\rangle \equiv \int d\,y_1 \dots \int d\,y_n\, P_n(1, \dots, n)\ y_1\, y_2 \dots y_n \tag{5.7}$$

A process is called stationary if all averages (or moments) are independent of the time origin so that it behaves the same way no matter when we start the measurements. This means every P_n is invariant if its time arguments are all displaced by the same amount:

$$P_n(y_1, t_1 \,;\, \dots \,;\, y_n, t_n) = P_n(y_1, t_1 + \tau \,;\, \dots \,;\, y_n, t_n + \tau) \qquad \forall\, n, \tau \tag{5.8}$$

Setting $\tau = -t_1$, we have

$$P_1(y_1, t_1) = P_1(y_1, 0) \equiv P_1(y_1) \tag{5.9}$$

so that $\langle\, y(t) \,\rangle$ is t-independent.

Similarly,

$$\begin{aligned} P_2\,(y_1, t_1; y_2, t_2) &= P_2\,(y_1, 0\,; y_2, t_2 - t_1) && (\ \tau = -t_1\ ) \\ &= P_2\,(y_1, t_1 - t_2; y_2, 0) && (\ \tau = -t_2\ ) \end{aligned}$$

so that

$$\begin{aligned} \langle\, y(t_1)\, y(t_2) \,\rangle &= \langle\, y(0)\, y(t_2 - t_1) \,\rangle \\ &= \langle\, y(t_1 - t_2)\, y(0) \,\rangle \end{aligned}$$

By definition, all processes in equilibrium are stationary.

Various P_n's can be related by conditional probabilities.

For example,

$$P_2(1, 2) = P_1(1)\, P_{1\,|\,1}(1 \mid 2) \tag{5.11}$$

where

$$P_{1|1}(1 \mid 2) \equiv P_{1|1}(y_1, t_1 \mid y_2, t_2)$$

is the conditional probability density that $Y = y_2$ at $t = t_2$ given that $Y = y_1$ at $t = t_1$.

Combining (5.5) & (5.11), we have

$$\begin{aligned} P_1(2) &= \int d\,y_1\, P_2(1, 2) \\ &= \int d\,y_1\, P_1(1)\, P_{1|1}(1 \mid 2) \qquad \forall\ t_1 \qquad (5.12) \end{aligned}$$

which relates P_1 of different arguments.

From (5.11), we have

$$\begin{aligned} \int d\,y_2\, P_{1|1}(1 \mid 2) &= \frac{1}{P_1(1)} \int d\,y_2\, P_2(1, 2) \\ &= \frac{1}{P_1(1)} P_1(1) \\ &= 1 \qquad \forall\ t_1\ \&\ t_2 \qquad (5.13) \end{aligned}$$

Joint conditional probability density of higher order can also be defined:

$P_{k|j}(1, \ \ldots, k \mid k+1, \ \ldots, k+j)$

$\equiv$ Joint conditional probability density that $Y = y_{k+1}$ at $t = t_{k+1}$, ..., and $Y = y_{k+j}$ at $t = t_{k+j}$, given that $Y = y_1$ at $t = t_1$, ..., and $Y = y_k$ at $t = t_k$. (5.14)

(5.11) thus generalizes to

$$\begin{aligned} & P_{k+j}(1, \ \ldots, k, k+1, \ \ldots, k+j) \\ &= P_k(1, \ \ldots, k)\, P_{k|j}(1, \ \ldots, k \mid k+1, \ \ldots, k+j) \qquad 1 \le k \qquad (5.15) \end{aligned}$$

Note that $P_{k|j}(1, \ \ldots, k \mid k+1, \ \ldots, k+j)$ is invariant under any permutations of the k (or j) arguments among themselves [c.f. (5.3a)].

Joint conditional probability densities describe memory effects. In this respect, it is convenient to assume that $t_1 < t_2 < \ldots < t_n$ for all n.

In which case, if there are no memory effects, then

$$P_{k|1}(1, \ \ldots, k \mid k+1) = P_1(k+1) \qquad \forall\ k \qquad (5.15a)$$

so that

$$\begin{aligned} & P_{k+1}(1, \ \ldots, k, k+1) \\ &= P_k(1, \ \ldots, k)\, P_1(k+1) \\ &= P_{k-1}(1, \ \ldots, k-1)\, P_1(k)\, P_1(k+1) \\ &\quad \vdots \\ &= P_1(1) \ldots P_1(k)\, P_1(k+1) \qquad (5.15b) \end{aligned}$$

More generally,

$$P_{k|j}(1, \ \ldots, k \mid k+1, \ \ldots, k+j) = P_j(k+1, \ \ldots, k+j) \qquad \forall\ k, j$$

so that (5.15) becomes

$$\begin{aligned} & P_{k+j}(1, \ \ldots, k, k+1, \ \ldots, k+j) \\ &= P_k(1, \ \ldots, k)\, P_j(k+1, \ \ldots, k+j) \end{aligned}$$

$$\vdots$$
$$= P_1(1) \ldots P_1(k)\, P_1(k+1) \ldots P_1(k+j)$$

A Markov process retains memory only of its immediate pass, i.e.,

$$P_{k|1}(1, \ \ldots, \ k \mid k+1) = P_{1|1}(k \mid k+1) \qquad \forall \ k \qquad (5.16)$$

so that

$$\begin{aligned} & P_{k+1}(1, \ \ldots, \ k, \ k+1) \\ &= P_k(1, \ \ldots, \ k)\, P_{1|1}(k \mid k+1) \\ &= P_{k-1}(1, \ \ldots, \ k-1)\, P_{1|1}(k-1 \mid k)\, P_{1|1}(k \mid k+1) \\ & \quad \vdots \\ &= P_1(1)\, P_{1|1}(1 \mid 2) \ \ldots P_{1|1}(k-1 \mid k)\, P_{1|1}(k \mid k+1) \qquad (5.17a) \end{aligned}$$

Comparing with (5.15b), $P_{1|1}(k \mid k+1)$ is called the transition probability.

Consider now the special case

$$P_3(1, 2, 3) = P_1(1)\, P_{1|1}(1 \mid 2)\, P_{1|1}(2 \mid 3) \qquad (5.17)$$

$$\rightarrow \quad P_2(1, 3) = \int d\, y_2\, P_3(1, 2, 3) \qquad \forall \ t_2 \qquad [\ (5.5) \text{ used }]$$

$$= P_1(1) \int d\, y_2\, P_{1|1}(1 \mid 2)\, P_{1|1}(2 \mid 3) \qquad (5.18)$$

On the other hand

$$P_2(1, 3) = P_1(1)\, P_{1|1}(1 \mid 3)$$

Comparing with (5.18) then gives

$$P_{1|1}(1 \mid 3) = \int d\, y_2\, P_{1|1}(1 \mid 2)\, P_{1|1}(2 \mid 3) \qquad \forall \ t_2 \qquad (5.19)$$

which is known as the Chapman-Kolmogorov equation.