§8. A NOISE PRIMER - MATHEMATICS

The theoretical treatment of noise and of random processes is an interesting and quite important subject in its own right, so before I clarify its physical basis and the way it enters mathematically into the description of physical systems, lets first recall the main features of noise and its mathematical description. The treatment will not be wide-ranging or comprehensive - we simply take what we need. I begin by recalling basic features of Probability theory, & probability distributions. Then we do the same for the theory of random processes (ie., different kinds of noise process). Finally, I briefly describe how to set up a path integral description of noise processes.

I do not discuss here disorder (ie., "noise in space"). Although some of the methods are similar, the physics is very different, particularly in quantum systems.

8(a) PROBABILITY DISTRIBUTIONS

Probability theory for ordinary random variables (as opposed to random processes) is concerned with calculating and using various probability distributions for these variables. Thus one defines a probability distribution $P(x)$ over a domain of possible "states", or values of x (which may be discrete or continuous, or both, and may have a complex topology). The probability that x lies in the range $x_0 \leq x \leq x_0 + dx$ is $P(x_0)\,dx$. All this is intuitively obvious, although a better formulation uses measure theory - one assigns, following the treatment the Kolmogorov school, a probability measure over the domain of possible states.

One convenient way of characterising $P(x)$ is by its moments. Another is by its cumulants, and another is by defining its characteristic function (which is basically just its Fourier transform). Looking at these in turn, we have

(i) MOMENTS We define the n-th moment of the probability distribution $P(x)$ as (assuming x is 1-dimensional):

$$\mu_n \equiv \langle x^n \rangle = \int dx\, x^n P(x) \qquad (8.1)$$

of which the most important examples are the 1st & 2nd moments:

$$\langle x \rangle =: \mu_1 = \int dx\, x P(x) \qquad \text{(MEAN)} \qquad (8.2)$$

$$\langle x^2 \rangle = \mu_2 = \int dx\, x^2 P(x) \qquad \text{(MEAN SQUARE)} \qquad (8.3)$$

but in physics one often uses the 3rd and 4th moments as well. For $P(x)$

to be a properly defined probability distribution one obviously requires that

$$\int dx\, P(x) = 1 \tag{8.4}$$

and also that

$$\langle f \rangle \equiv \langle f(x) \rangle = \int dx\, f(x) P(x) \tag{8.5}$$

be the expectation value of any function of x defined over its domain. Obviously one can also characterise $P(x)$ completely using any set of orthonormal functions defined over this domain - the most common is the

(ii) <u>CHARACTERISTIC FUNCTION</u> which is just the Fourier transform of $P(x)$; we shall simply write it as

$$\mathcal{P}(k) = \langle e^{ikx} \rangle = \int dx\, e^{ikx} P(x) \tag{8.6}$$

with Fourier transform

$$P(x) = \int dk\, e^{-ikx} \mathcal{P}(k) \tag{8.7}$$

and from which the moments are immediately extracted using:

$$\left.\begin{aligned} \mathcal{P}(k) &= \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \mu_n \\ \text{ie.,} \quad \mu_n &= (-i)^n \frac{d^n}{dk^n} \mathcal{P}(k) \Big|_{k=0}. \end{aligned}\right\} \tag{8.8}$$

Another function often used is the "probability generating fn." $F(z) = \langle z^x \rangle$, but we will not use this here. Another common characterisation is via the

(iii) <u>CUMULANTS</u> : These are most easily defined via the Taylor expansion of $\log \mathcal{P}(k)$, as

$$\log \mathcal{P}(k) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} c_n \tag{8.9}$$

so that the c_n are related to the μ_n by comparing (8.9) & (8.8); one gets

$$\left.\begin{aligned} c_1 &= \mu_1 \\ c_2 &= \mu_2 - \mu_1^2 \\ c_3 &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 \\ c_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4 \\ &\text{etc.} \end{aligned}\right\} \tag{8.10}$$

The 2nd cumulant is very well-known:

$$c_2 = \mu_2 - \mu_1^2 \equiv \langle (x - \langle x \rangle)^2 \rangle = \sigma^2 \quad \text{(VARIANCE)} \tag{8.11}$$

where σ is the "standard deviation". Cumulants are very useful in both stat. mech. and in field theory, because one takes logs of distribution functions or field distributions to get physical quantities.

All of this is very simple, but what is much more interesting is when we have several different values, and we are interested in correlations between them. We go through the same routine as before. The probability distribution is $P(\{x_j\}) \equiv P(\underline{X}) \equiv P(x_1, x_2 \ldots x_m)$, satisfying.

$$\int d\underline{x}\, P(\underline{x}) = \int dx_1 \ldots dx_m\, P(x_1 \ldots x_m) = 1 \qquad (8.12)$$

$$\int d\underline{x}\, f(\underline{x}) P(\underline{x}) = \langle f(\underline{x}) \rangle \qquad (8.13)$$

and now having a large variety of moments:

$$\langle x_1^{n_1} x_2^{n_2} \ldots x_m^{n_m} \rangle = \int d\underline{x}\; x_1^{n_1} x_2^{n_2} \ldots x_m^{n_m} P(\underline{x}) = \underset{\sim}{\mu}_m (x_1^{n_1}, x_2^{n_2} \ldots) \qquad (8.14)$$

The characteristic fn. and cumulants are defined similarly:

$$\left.\begin{aligned} P(\underline{k}) = P(k_1 \ldots k_m) &= \langle e^{i\underline{k}\cdot\underline{x}} \rangle \equiv \langle e^{i(k_1 x_1 + \ldots k_m x_m)} \rangle \\ &= \sum_{n_1} \sum_{n_2} \ldots \sum_{n_m} \frac{(ik_1)^{n_1} (ik_2)^{n_2} \ldots (ik_m)^{n_m}}{n_1!\, n_2! \ldots n_m!} \langle x_1^{n_1} x_2^{n_2} \ldots x_m^{n_m} \rangle . \end{aligned}\right\} \qquad (8.15)$$

$$\ln P(\underline{k}) \equiv \sum_{n_1} \ldots \sum_{n_m} \frac{(ik_1)^{n_1} (ik_2)^{n_2} \ldots (ik_m)^{n_m}}{n_1!\, n_2! \ldots n_m!} \underset{\sim}{\chi}_m (x_1^{n_1}, x_2^{n_2} \ldots x_m^{n_m}) \qquad (8.16)$$

What is of most interest in a multivariate probability distribution is the correlation between different variables - for this reason the cumulants are often called "correlation functions". The most important of these is the (normalised) "covariance", defined by

$$\begin{aligned} \mathrm{cov}(x_n, x_m) \equiv \bar{X}_2 (x_n, x_m) &= \frac{\underset{\sim}{\chi}_2 (x_n, x_m)}{(\langle C_n^2 \rangle \langle C_m^2 \rangle)^{1/2}} \\ &\equiv \frac{\langle (x_1 - \langle x_1 \rangle)(x_2 - \langle x_2 \rangle) \rangle}{(\langle C_n^2 \rangle \langle C_m^2 \rangle)^{1/2}} \\ &\equiv \frac{\langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle}{\left[(\langle x_n^2 \rangle - \langle x_n \rangle^2)(\langle x_m^2 \rangle - \langle x_m \rangle^2) \right]^{1/2}} \end{aligned} \qquad (8.17)$$

We are of course all familiar with examples of ways in which random variables may be correlated (or anti-correlated), or be statistically independent. Thus it seems natural to believe the following propositions:

$$\left.\begin{array}{l} \text{(i) The results of tossing 2 different coins are uncorrelated - or indeed} \\ \quad\;\; \text{of tossing } m \text{ different coins.} \\[1em] \text{(ii) The atmospheric pressure, in a certain location, and the precipitation in} \\ \quad\;\; \text{the same location, are correlated.} \end{array}\right\} \qquad (8.18)$$

We actually have very strong beliefs, whose basis I will not analyze here, concerning the significance of correlations or lack thereof, between random variables – thus, eg., the lack of any causal connection between 2 variables is assumed to lead to their statistical independence (and a correlation is also taken as evidence for, although certainly not proof, of some causal or logical connection between them).

It is noteworthy that many physicists tend to forget these beliefs, and the reasoning behind them, when it comes to quantum phenomena.

CHANGING VARIABLES IN PROBABILITY DISTRIBUTIONS

Suppose we know the probability distribution over one variables x, but we are really interested in the distribution over another variable z, which is a function $z = f(x)$. Then the probability distribution over z is just

$$\hat{P}(z) = \int dx\, P(x)\, \delta(z - f(x)) \qquad (8.19)$$

(this can indeed be thought of as a def^{n} of $P(z)$); the generalisation to multivariate distributions is obvious. One common transformation of variables arises when one is interested in the distribution over sums, differences, etc., of random variables. For example, suppose we know a distribution $P(x_1, x_2)$ for 2 variables, and we want to know the distribution $\hat{P}(z)$, where $z = A_1 x_1 + A_2 x_2$. It is then obvious that this is given by:

$$z = A_1 x_1 + A_2 x_2 \;; \qquad P(z) = \int dx_1 dx_2\, P(x_1, x_2)\, \delta(A_1 x_1 + A_2 x_2 - z) \qquad (8.20)$$

and more generally if

$$z = \sum_{r=1}^{m} a_r x_r^{n_r} \quad \Rightarrow \quad \bar{P}(z) = \prod_{r=1}^{m} \int dx_r\, P(x_1, \ldots x_m)\, \delta(z - \sum_r x_r^{n_r}) \qquad (8.21)$$

The characteristic function $\hat{P}_q(q)$ of $\hat{P}(z)$ is given simply by

$$P(q) = \langle e^{iqz} \rangle = \langle e^{iqf(x)} \rangle = \int dx\, e^{iqf(x)} \qquad (8.22)$$

8(b) EXAMPLES OF PROBABILITY DISTRIBUTIONS

Here we just very briefly recall the common probability distributions, and the relationships between them – I assume you know them all already. Perhaps the most basic is the

(i) BINOMIAL DISTRIBUTION: We imagine a random variable which can take only 2 values (eg., ± 1), with probabilities p_+, p_-, such that $p_+ + p_- = 1$. We now ask first what is the probability of getting, in N trials, the sequence $(x_1, x_2, \ldots x_N)$, in which we have n_+ results with $+$, and n_- results with $-$. The answer for a GIVEN SEQUENCE is

$$P(x_1 \ldots x_N)\Big|_{\substack{n_+ \\ n_-}} = p_+^{n_+} p_-^{n_-} \qquad (8.23)$$

On the other hand the probability of getting n_+ and n_- results for $\pm$, in any order (ie., summed over all possible sequences with these numbers) is just

$$\left.\begin{aligned} p_N(n_+,n_-) &= \frac{N!}{n_+!\,n_-!}\, p^{n_+} p_-^{n_-} \\ &\equiv \frac{N!}{(N-n_+)!\,n_+!}\, p_+^{n_+}(1-p_+)^{N-n_+} \end{aligned}\right\} \qquad (8.24)$$

where the binomial coefficient

$$C^N_{n_+} = \frac{N!}{(N-n_+)!\,n_+!} \qquad (8.25)$$

is just the total number of sequences satisfying this condition. Standard results for the binomial distribution are

$$\langle n_+\rangle = \sum_{n_+} n_+ p_N(n_+,n_-) = p_+\frac{\partial}{\partial p_+}\sum_{n_+} p_N(n_+,n_-)$$

ie.
$$\mu_1 \equiv \langle n_+\rangle = Np_+ \qquad \text{(MEAN)} \qquad (8.26)$$

$$\begin{aligned} \mu_2 &= \sum_{n_+} n_+^2 p_N(n_+,n_-) \\ &= (Np_+)^2 + Np_+(1-p_+) \qquad \text{(MEAN SQUARE)} \qquad (8.27) \end{aligned}$$

$$\left.\begin{aligned} \sigma_N^2 = \mu_2-\mu_1^2 &= Np_+(1-p_+) \\ \sigma_N &= N^{1/2}(p_+(1-p_+))^{1/2} \end{aligned}\right\} \text{VARIANCE}$$

The meaning of the binomial distribution is obvious – you are all familiar with examples, ranging from toin cosses and poker hands to balls in urns or atoms in compartments (with the necessary generalisation to multinomial distributions in some cases). Now consider that most of the distributions you will come across are derivable from it:

(ii) <u>GAUSSIAN DISTRIBUTION</u> : This can be derived from the binomial distribution in the limit where $N \gg 1$ and $Np_+ \gg 1$. The derivation is found in many texts; one uses Stirling's formula for the factorial, viz.,

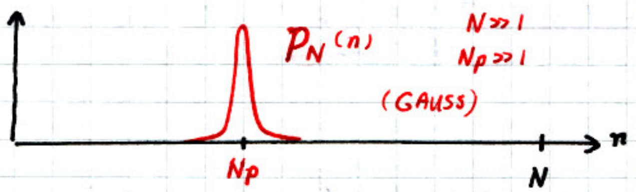
$$n! \sim (2\pi n)^{1/2}\,(n/e)^n \qquad (8.28)$$

to expand $p_N(n)$ (where $n_+ \to n$). One then expands the result in a Taylor expansion about $x = x_0$, where $x = n/N$ and $x_0 = \mu_1/N = p$ (where now $p \to p_+$) are treated as continuous variables in the large N limit. It is then found that one gets, to a very good approximation, the standard result for a Gaussian distribution:

$$\mathcal{P}(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-(x-x_0)^2/2\sigma^2} \qquad (8.27)$$

where $\quad \sigma^2 = p(1-p)/N \qquad x_0 = p \qquad (8.28)$

is the value for the variance that is obtained in this derivation, around the mean x_0; The variance σ_N^2 of the initial distribution (not in terms of the renormalised variable $x = n/N$) is



$$\sigma_N^2 = Np(1-p) \qquad (8.29)$$

and the mean of this initial distribution $p_N(n)$ is $\mu_1 = Np$.

The binomial derivation is just a prop for the Gaussian - it is much more generally applicable, as we shall see below. Thus we can simply drop all reference to p and N, and work exclusively with $\mathcal{P}(x)$ in the form in (8.27). Note that the characteristic function of $\mathcal{P}(x)$ is just

$$\mathcal{P}(k) = \int dx\, e^{ikx} \mathcal{P}(x) = e^{ikx_0 - \frac{1}{2}k^2\sigma^2} \qquad (8.30)$$

from which we see that only the 1st & 2nd cumulants are non-zero, ie.,

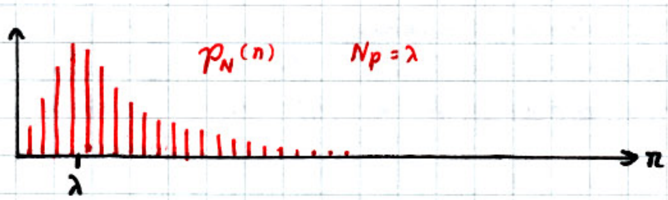
$$c_1 = x_0 \qquad c_3, c_4, \ldots = 0 \qquad \text{(GAUSSIAN DIST.)} \qquad (8.31)$$
$$c_2 = \sigma^2$$

(iii) <u>POISSON DISTRIBUTION</u> : Again, we can get this from the binomial distribution, but this time we assume $N \gg 1$, $p \ll 1$, and $Np = \lambda$; and we assume that as $N \to \infty$, λ is a constant (so that $p \to 0$). It then follows that

$$\frac{N!}{(N-n)!\,n!} \sim \frac{N^n}{n!} \qquad \text{and} \qquad (1-p)^{N-n} \sim (1-p)^N \qquad (8.32)$$
$$\equiv e^{-\lambda}$$

using $e^x = \lim_{N\to\infty} (1 + x/N)^N$. Thus the distribution tends to

$$\mathcal{P}_N(n) \to \frac{\lambda^n e^{-\lambda}}{n!} \qquad \text{(POISSON)} \qquad (8.33)$$



for which one has

$$\mu_1 = \sigma^2 = \lambda \qquad (8.34)$$

Just like the Gaussian distribution, the Poisson distribution arises in many contexts, not just as a limiting case of a Gaussian distribution. And it is trivial to give multivariate generalisations of them, to a set of variables each taking values $\pm$, or to a single variable which ranges over, say S different values, ...

The more general application of the Gaussian distribution is basically a consequence of the following result. Suppose we consider a random variable x, and then make N measurements of it; or alternatively, we consider N different random variables, each with the probability distribution $P(x)$, and make one measurement on each. Then the mean will be given by the random variable

$$y_N = \frac{1}{N} \sum_{j=1}^{N} x_j \tag{8.35}$$

and the obvious question is – what is the distribution $\tilde{P}(y)$ for y? The famous "central limit theorem" states that this distribution is

$$\tilde{P}(y) \approx \left(\frac{N}{2\pi\sigma^2}\right)^{\frac{1}{2}} e^{-\frac{N}{2\sigma^2}(y-\langle x\rangle)^2} \quad \text{(CENTRAL LIMIT)} \tag{8.36}$$

where σ is basically the variance of $P(x)$. The proof assumes that (i) N is large, that (ii) the $\{x_j\}$ are statistically independent, and (iii) that the moments of $P(x)$ are well-behaved (in particular, that they are FINITE). Let us consider the characteristic function

$$\tilde{P}(q) = \int dy\, e^{iq(y-\langle x\rangle)} \tilde{P}(y-\langle x\rangle) = \prod_{j=1}^{N} \int dx_j\, P(x_j) e^{i\frac{k}{N}(x_j - \langle x\rangle)}$$

$$= \left(P(k/N)\right)^N \tag{8.37}$$

where

$$P(k/N) = \int dx\, P(x) e^{i\frac{k}{N}(x-\langle x\rangle)} \tag{8.38}$$

$$= 1 - \tfrac{1}{2} \tfrac{k^2}{N^2} \sigma^2 + \cdots \tag{8.39}$$

and, as usual, $\sigma^2 = \langle x^2\rangle - \langle x\rangle^2$. Now, the argument goes, the effect of phase oscillations in (8.38) is to make $P(k/N)$ decrease with larger k, and unless some of the higher moments in $P(x)$ are very large, then as N becomes very large, the N-th power of $P(k/N)$ decreases very fast with k, so that we can write

$$\tilde{P}(q) \sim \left(1 - \frac{k^2\sigma^2}{2N^2} + \ldots\right)^N$$

$$\xrightarrow[N\to\infty]{} e^{-k^2\sigma^2/2N} \tag{8.40}$$

and then (8.36) follows immediately by Fourier transforming.

Thus, physically, Gaussian distributions are frequent when a large number of independent degrees of freedom are present; or a large number of trials or events; etc...

The Poisson distribution also arises very generally simply because it gives the distribution in time (or some other "space") of a set of randomly distributed (uncorrelated) events (the distribution in time of decay events of a set of, e.g., radioactive atoms; the sound of raindrops on a roof, etc... BUT NOT, e.g., the time of passage of cars on an autoroute), in which the mean number per unit time is λ.

There are of course many many other useful and commonly occurring distributions. In physics, for example, we often use the Lorentz distribution:

$$P(x) = \frac{1}{\pi} \frac{\Gamma}{(x-x_0)^2 + \Gamma^2} \tag{8.41}$$

which is actually rather pathological, since it has long "Lorentz wings" and infinite moments - which are usually cut off at some "UV cut-off".

This is all we will need for the moment, concerning elementary probability theory.

8(c) PROBABILITY DISTRIBUTIONS OVER CLASSES OF FUNCTIONS

What we are interested in, in both the formulation of noise theory and, indeed, the formulation of field theory (quantum or classical) is the probability of a particular FUNCTION $f(x)$ occurring, amongst some large (but usually restricted) SET of functions. Thus we need now to assign some probability, not to a variable x, but to a function $f(x)$; and this probability $P[f(x)]$ is a functional defined over the relevant class of functions. This notion is one you are already used to - for example, in the idea that the energy E of a system is a functional $E[n(\underline{r},t)]$ of the density field $n(\underline{r},t)$ of a system, or in the idea that the amplitude $G(Q_2, Q_1; t)$ for a quantum system to go from Q_1 at time $t=0$, to Q_2 at time t, is a functional $G[q(\tau)]$ of all the possible paths between.

For the particular case of noise processes, restrictions on the class of possible functions are usually imposed by the physicist, starting from various assumptions discussed in the next section (§9). Here I give an elementary introduction to the mathematics.

Let us first consider how we can characterise a noise process, as a particular kind of problem amongst the more general set of questions arising when one tries to characterise an infinite set of functions (a very large set in general, in fact of size).

We can of course do this by defining a probability functional $P[\xi(t)]$ over a set of functions $\xi(t)$, themselves functions of time. What then specifies WHAT KIND of functions are allowed? Here are 3 ways of doing this, similar to those discussed for ordinary probabilities above:

(i) <u>MOMENTS</u>: The moment functionals are an obvious generalisation of the moments defined for a multivariate probability function (recall eqtn (8.14)). We define a set of functions

$$M_n^\xi(t_1 \dots t_n) = \langle \xi(t_1)\,\xi(t_2)\dots\xi(t_n)\rangle \tag{8.42}$$

ie., the infinite sequence $\quad M_1^\xi(t_1),\ M_2^\xi(t_1,t_2),\dots \quad (8.43)$

which, continued to the dense set of times $\{t_1, t_2, \dots t_n, \dots\}$ gives a complete description of the random process $\xi(t)$. These moments are defined as

$$\left.\begin{aligned} M_1^\xi(t_1) &= \int \mathcal{D}\xi(t)\, P[\xi(t)]\ \xi(t=t_1) \\ M_2^\xi(t_1,t_2) &= \int \mathcal{D}\xi(t)\, P[\xi(t)]\ \xi(t=t_1)\,\xi(t=t_2). \end{aligned}\right\} \tag{8.44}$$

etc

but I emphasize that this is not necessarily a definition – the hierarchy of moments in (8.43) by itself gives us a definition of $P[\xi(t)]$.

(ii) CHARACTERISTIC FUNCTIONAL : Supposing instead of the hierarchy of moments we enumerated the following sequence of "characteristic functions" or expectation values:

$$\left.\begin{aligned} w_1(u_1;t_1) &= \langle e^{iu_1\xi(t_1)}\rangle. \\ w_2(u_1u_2;t_1t_2) &= \langle e^{i(u_1\xi(t_1)+u_2\xi(t_2))}\rangle \\ w_n(u_1..u_n;t_1..t_n) &= \langle e^{i\sum_{r=1}^n u_r\xi(t_r)}\rangle \quad ; \text{ etc.} \end{aligned}\right\} \tag{8.45}$$

which can be expanded as

$$\begin{aligned} w_n(u_1\dots u_n;t_1,\dots t_n) &= \langle e^{i\sum_{r=1}^n u_r\ \xi(t_r)}\rangle \\ &\equiv 1+\sum_{l=1}^\infty \frac{i^l}{l!}\left\{\prod_{j=1}^{l}\sum_{\alpha_j=1}^{n} M_l(t_{\alpha_1}\dots t_{\alpha_l})\,u_{\alpha_1}\dots u_{\alpha_l}\right\} \end{aligned} \tag{8.46}$$

If you find this last form complicated, note that from (8.45) we have

$$w_1(u_1,t_1) = \langle e^{iu_1\xi(t_1)}\rangle = 1 + iu_1\langle\xi(t_1)\rangle - \tfrac{1}{2}u_1^2\langle\xi^2(t_1)\rangle + \dots \tag{8.47}$$

$$\begin{aligned} w_2(u_1u_2;t_1t_2) &= \langle e^{i(u_1\xi(t_1)+u_2\xi(t_2))}\rangle \\ &\equiv 1 + i\big(u_1\langle\xi(t_1)\rangle + u_2\langle\xi(t_2)\rangle\big) - \tfrac{1}{2}\Big(u_1^2\langle\xi^2(t_1)\rangle + u_2\langle\xi^2(t_2)\rangle \\ &\qquad + u_1u_2\big(\langle\xi(t_1)\xi(t_2)\rangle\big)^2\Big) + \dots \\ &\text{etc.} \end{aligned} \tag{8.48}$$

which is what (8.46) is saying. Now, if we go to the limit $n\to\infty$, in such a way that the $\{t_{\alpha_j}\}$ cover the time domain densely, we have the CHARACTERISTIC FUNCTIONAL

$$\hat{P}[u(t)] = \langle e^{i\int dt\, u(t)\xi(t)} \rangle \tag{8.49}$$

$$\equiv \frac{\int \mathcal{D}\xi(t)\, P[\xi(t)]\, e^{i\int dt\, u(t)\xi(t)}}{\int \mathcal{D}\xi(t)\, P[\xi(t)]} \tag{8.50}$$

where we have assumed that $P[\xi(t)]$ may not be normalised (which was not actually enforced in (8.44) but is assumed). Note that the moments in (8.44) are easily obtained from the general result

$$\langle \mathcal{F}[\xi(t)] \rangle = \frac{\int \mathcal{D}\xi(t)\, P[\xi(t)]\, \mathcal{F}[\xi(t)]}{\int \mathcal{D}\xi(t)\, P[\xi(t)]} \tag{8.51}$$

for the average of a Functional $\mathcal{F}$ of $\xi(t)$, by assuming that $\mathcal{F}$ takes the simple form

$$\left.\begin{aligned} \mathcal{F}[\xi] &= e^{i\int dt\, u(t)\xi(t)} \\ u(t) &= \sum_{\alpha_j} u_{\alpha_j}\, \delta(t - t_{\alpha_j}) \end{aligned}\right\} \text{ gives moments} \tag{8.52}$$

(compare (8.44) and (8.46)).

The inverse transformation to (8.49) is of course

$$P[\xi(t)] = \int \mathcal{D}u(t)\, \hat{P}[\xi(t)]\, e^{-i\int dt\, u(t)\xi(t)} \tag{8.53}$$

where no normalisation factor is now necessary. Let us also note that if we apply the inverse transformation to the set of characteristic functions in (8.45) or (8.46), we get

$$W_n(\xi(t_1)\xi(t_2)\ldots\xi(t_n)) = \prod_{j=1}^{n} \int_{-\infty}^{\infty} \frac{du_j}{2\pi}\, e^{i\sum_{j=1}^{n} u_j \xi(t_j)}\, \varpi(u_1,\ldots u_n;\, t_1,\ldots t_n) \tag{8.54}$$

where the set $\{W_n\}$ constitute yet another way (indeed maybe the most natural way) of characterising the probability functional $P[\xi(t)]$; the function $W_n(\xi(t_1)\ldots\xi(t_n))$ is the probability density for the set of random variables $\xi(t_1),\ldots \xi(t_n)$; as an average it is just

$$W_n(z_1,\ldots z_n;\, t_1\ldots t_n) = \langle \delta(z_1 - \xi(t_1))\, \delta(z_2 - \xi(t_2)) \cdots \delta(z_n - \xi(t_n)) \rangle. \tag{8.55}$$

(iii) CORRELATION FUNCTIONS : In a fairly obvious generalisation of what we did for cumulants of probability functions, we write for the logarithm of the characteristic functional that

$$\log \tilde{P}[u(t)] = \sum_{l=1}^{\infty} \frac{i^l}{l!} \prod_{r=1}^{l} \int dt_r \; K_r(t_1 \ldots t_r)\, u(t_r)$$

$$\equiv \sum_{l=1}^{\infty} \frac{i^l}{l!} \iiint dt_1 \ldots dt_l \; K_l(t_1 \ldots t_l)\, u(t_1) \cdots u(t_l) \qquad (8.56)$$

where the correlation functions

$$K_l(t_1 \ldots t_l) \equiv K[\xi(t_1) \ldots \xi(t_l)] \qquad (8.57)$$

are related to the moments in a way analogous to the relation (8.10) between moments and cumulants:

$$M_1(t_1) = K_1(t_1) \qquad (8.58)$$

$$M_2(t_1 t_2) = K_2(t_1 t_2) + K_1(t_1) K_1(t_2) \qquad (8.59)$$

$$M_3(t_1 t_2 t_3) = K_3(t_1 t_2 t_3) + \big(K_1(t_1) K_2(t_2 t_3) + K_1(t_2) K_2(t_1 t_3) + K_1(t_3) K_2(t_1 t_2)\big) + K_1(t_1) K_1(t_2) K_1(t_3) \qquad (8.60)$$

etc.

with the inverse transformation (ie., K_3 in terms of the M_l) obvious from eqtn (8.10). Finally, we note that the relationship to the characteristic function in (8.45) and (8.46) is

$$\log w_n(u_1 \ldots u_n ; t_1 \ldots t_n) = \sum_{l=1} \frac{i^l}{l!} \left\{ \prod_{j=1}^{l} \sum_{\alpha_j=1}^{n} K_l(t_{\alpha_1} \ldots t_{\alpha_l})\, u_{\alpha_1} \ldots u_{\alpha_l} \right\} \qquad (8.61)$$

8(d) GAUSSIAN RANDOM PROCESSES

The analogue of a Gaussian distribution for a random variable, when we go to random functions, is a "Gaussian random process" – for random processes occurring in time this is "Gaussian noise". There are various ways to approach it formally – here I discuss its ~~own~~ mathematical origin in

(1) HEURISTIC DERIVATIONS : The following derivation are meant to show the relationship to the discussion of Gaussian probabilities resulting from a sum over many

random variables (eqtns (8.35) - (8.40)), and also to give you a more intuitive feeling for what is going on than what the mere mathematics can provide.

Let's begin with the specification of a Gaussian noise process. We call a random process $\xi(t)$ Gaussian if the probability density functional is

$$\mathcal{P}_G[\xi(t)] = e^{-\frac{1}{2}\int dt \int dt' \xi(t) Q(t-t') \xi(t')} \qquad \begin{pmatrix}\text{Gaussian,} \\ \text{zero mean}\end{pmatrix} \qquad (8.62)$$

so that a variable $\mathcal{F}[\xi(t)]$ has, from (8.51), the expectation value

$$\langle \mathcal{F}_G[\xi] \rangle = \frac{\int \mathcal{D}\xi(t) \mathcal{P}_G[\xi] \mathcal{F}[\xi]}{\int \mathcal{D}\xi(t) \mathcal{P}_G[\xi]}$$

$$= \frac{1}{\|Q(t-t')\|^{1/2}} \int \mathcal{D}\xi(t) \mathcal{F}[\xi] e^{-\frac{1}{2}\int d\tau \int d\tau' \xi(\tau) Q(\tau-\tau') \xi(\tau')} \qquad (8.63)$$

where we shall show below that the <u>determinant</u> of the "operator" $Q(t-t')$ normalises the probability density according to (8.23). The distribution (8.63) has zero mean, ie.,

$$\langle \xi(t) \rangle = 0 \qquad (8.64)$$

and we will also see that $\quad \langle \xi(t) \xi(t') \rangle = \mathcal{X}(t-t') \qquad (8.65)$

with the definition. $\quad \int dt' \, \mathcal{X}(t_1 - t') \, Q(t' - t_2) = \delta(t_1 - t_2) \quad (8.66).$

All higher moments / correlation functions of $\xi(t)$ are given in terms of the basic function $Q(t-t')$. Finally, we shall see that the characteristic functional of $\mathcal{P}[\xi(t)]$ is given by

$$\mathcal{P}_G[u(t)] = \langle e^{i\int dt\, u(t) \xi(t)} \rangle$$

$$\equiv \frac{1}{\|Q(t-t')\|^{1/2}} e^{-\frac{1}{2}\int dt \int dt'\, u(t) \mathcal{X}(t-t') u(t')} \qquad (8.67).$$

Just by looking at (8.62) we see roughly what it is describing – the noise process is such that $\xi(t)$ and $\xi(t')$ are mutually constrained in all "regions" where $Q(t-t')$ is large. Thus, eg., if $Q(t-t')$ is very large for times such that $|t-t'| > \tau_0$, then it is clear that in these regions of time, the product $\xi(t)\xi(t')$ should remain small on average. We shall see more of what Gaussian noise characteristics follow from (8.62) in a moment. Note incidentally, that the def^n of $\chi(t-t')$ in (8.66) in terms of $Q(t-t')$ as an "inverse operator" is most easily implemented in a calculation by Fourier transforming:

$$\int dt'\, Q(t_1-t')\,\chi(t'-t_2) \quad \Longrightarrow \quad Q(\omega)\,\chi(\omega) = 1. \tag{8.68}$$

where $$Q(\omega) = \int dt\, e^{i\omega t}\, Q(t) \tag{8.69}$$

and so we have
$$\chi(t) = \int \frac{d\omega}{2\pi}\, \frac{1}{Q(\omega)}\, e^{-i\omega t} \tag{8.70}$$

Now, let's see in a heuristic way how the results (8.63)-(8.67) are obtained. To do this, we will treat all the functional integrals as analogous to multiple integrals, extrapolating in a non-rigourous way finite-dimensional integrals to infinite dimensions. First let's recall the following simple results for Gaussian integrals:

$$\left.\begin{aligned} \int_{-\infty}^{\infty} \frac{dx}{(2\pi)^{1/2}}\, e^{-(\frac{1}{2}Qx^2 + bx + c)} &= \frac{1}{\sqrt{Q}}\, e^{b^2/2Q - c} \\ \int_{-\infty}^{\infty} \frac{dx}{(2\pi)^{1/2}}\, e^{-(\frac{1}{2}Qx^2 + ikx)} &= \frac{1}{\sqrt{Q}}\, e^{-k^2/2Q} \end{aligned}\right\} \tag{8.71}$$

so that
$$\left.\begin{aligned} \int \frac{d^n x}{(2\pi)^{n/2}}\, e^{-(\frac{1}{2}x_i Q_{ij} x_j + b_j x_j)} &= \frac{1}{|Q_{ij}|^{1/2}}\, e^{\frac{1}{2} b_i Q^{-1}_{ij} b_j} \\ \int \frac{d^n x}{(2\pi)^{n/2}}\, e^{-(\frac{1}{2}x_i Q_{ij} x_j + ik_j x_j)} &= \frac{1}{|Q_{ij}|^{1/2}}\, e^{-\frac{1}{2} k_i Q^{-1}_{ij} k_j} \end{aligned}\right\} \tag{8.72}$$

where $|Q_{ij}| = \det Q_{ij}$ is the determinant of the matrix and $\underline{x} = (x_1, x_2, .. x_n)$; and the inverse matrix $\chi_{ij} \equiv Q^{-1}_{ij}$ is defined by
$$Q^{-1}_{ik} Q_{kj} \equiv Q_{ik} Q^{-1}_{kj} = \delta_{ij} \tag{8.73}$$

Then, to define a functional integral over a Gaussian functional, we simply divide the timeline into N intervals, let $N \to \infty$, and we get the generalisations.

$$\int \frac{d^n x}{(2\pi)^{n/2}} \quad \to \quad \lim_{N\to\infty} \prod_{j=1}^{N} \int \frac{d\xi(t_j)}{(2\pi)^{n/2}} \quad \to \quad \int \mathcal{D}\xi(t) \tag{8.74}$$

$$e^{-\frac{1}{2}x_i Q_{ij} x_j + ik_j x_j} \quad \to \quad \exp\left\{-\tfrac{1}{2}\int dt \int dt'\, \xi(t)\, Q(t-t')\, \xi(t') + i\int dt\, u(t)\,\xi(t)\right\} \tag{8.75}$$

and then we easily that (8.67) is the generalisation of (8.72), and that (8.66) or (8.68) are just the generalisation of (8.73). The results (8.64) and (8.65) are produced by obvious manoeuvres as well. Thus, we have

$$\begin{aligned} \langle \xi(t) \rangle &= -i \frac{\delta}{\delta u(t)} P[u] \Big|_{u=0} \\ &= i \left[\int dt' \, \chi(t-t') u(t') \, P[u] \right] \Big|_{u(t)=0} = 0 \end{aligned} \tag{8.76}$$

by an obvious generalisation of the usual result that $-i \, d/dk \int dx \, e^{-\frac{1}{2}Qx^2 + ikx}$ is just ik/Q

It will be clear from above that it is often simpler to look at all of this in frequency space. With the same convention for Fourier transforms as in (8.69), we have

$$\left. \begin{aligned} P_G[\xi(\omega)] &= e^{-\frac{1}{2}\int \frac{d\omega}{2\pi} \Phi(\omega) |\xi(\omega)|^2} \\ P_G[u(\omega)] &= e^{-\frac{1}{2}\int \frac{d\omega}{2\pi} \chi(\omega) |u(\omega)|^2} \, |\chi(\omega)|^{1/2} \end{aligned} \right\} \tag{8.77}$$

where now we also notice that the correlation function $\chi(\omega)$ also defines the POWER SPECTRUM of the noise, ie.,

$$P(\omega) \equiv \langle |\xi(\omega)|^2 \rangle = \chi(\omega) \tag{8.78}$$

A special limiting case of a Gaussian random process is one in which both $\Phi(t-t')$ and $\chi(t-t')$ are "delta-correlated", ie., we have

$$\left. \begin{aligned} \chi(t-t') &= P_0 \, \delta(t-t') \quad &\text{so} \quad \chi(\omega) &= P_0 \\ \Phi(t-t') &= 1/P_0 \, \delta(t-t') \quad &\text{"} \quad \Phi(\omega) &= 1/P_0 \end{aligned} \right\} \begin{pmatrix} \text{WHITE} \\ \text{NOISE} \end{pmatrix} \tag{8.79}$$

For such a "white noise" spectrum, the probability distribution and the characteristic functionals are

$$\left. \begin{aligned} P_G[\xi] &= e^{-\frac{1}{2P_0}\int dt \, \xi^2(t)} = e^{-\frac{1}{2P_0}\int \frac{d\omega}{2\pi} |\xi(\omega)|^2} \\ P_G[u] &= P_0^{1/2} \, e^{-P_0/2 \int dt \, u^2(t)} = P_0^{1/2} \, e^{-P_0/2 \int \frac{d\omega}{2\pi} |\xi(\omega)|^2} \end{aligned} \right\} \tag{8.80}$$

Physically, one uses this form when the power spectrum of the noise is spread accross a very wide frequency range, extending well above and below the typical frequency response of the system we are interested in; and the power spectrum is found to be relatively flat throughout this frequency range, of magnitude P_0. One should not take the delta-function too literally – any real physical noise cannot fluctuate infinitely rapidly.

Finally, in this discussion of Gaussian random processes, note that if the mean value $\langle \xi(t) \rangle$ is finite, then the results in (8-62) et seq. generalise very simply. We have

$$P_G[\xi(t)] = e^{-\frac{1}{2}\int dt \int dt' \, (\xi(t)-\xi_0(t))\, \Phi(t-t')\, (\xi(t')-\xi_0(t'))}$$

$$P_G[u(t)] = \frac{1}{\| \Phi(t-t') \|} \, e^{i\int dt \, u(t)\xi_0(t)} \, e^{-\frac{1}{2}\int dt \int dt' \, u(t)\chi(t-t')u(t')}$$

where $\xi_0(t)$ is a specified function around which $\xi(t)$ is fluctuating (ie., we do NOT average over $\xi_0(t)$). This result would apply, eg., to a system in which the noise fluctuated around some very slowly-varying but known function $\xi_0(t)$.

(2) SOME ILLUSTRATIVE EXAMPLES :

Examples of processes in Nature which are well-described as Gaussian are actually quite common.