ERRATA

IRREVERSIBLE STATISTICAL MECHANICS

by

John M. Dawson
Princeton University

Course 554
Department of Astrophysical Sciences
February 10, 1967

On pages 48 through 52, note that equation numbers 190 through 198 have been used for two different sets of equations.

REVIEW OF SOME BASIC IDEAS FROM STATISTICAL MECHANICS

John M. Dawson

Probability

Consider that there is some physical situation or process which we are interested in. Let ξ be some quantity associated with this situation. Some examples are

1) Flipping coins ξ = heads or tails

2) Rolling a dice $\xi = 1, 2, \ldots, 6$

3) Number of molecules in a small volume $\xi = 1, 2, \ldots, N$

N is the maximum possible number of particles which can be put in the volume.

4) Velocity of a particle in the x-direction $-\infty < \xi < \infty$

5) Amplitude of the x-component of the electric field at a point in a plasma $-\infty < \xi < \infty$.

Depending on what we are investigating, x may be a discrete or a continuous variable. To define probability we must imagine that we have many examples of the situation of interest, ideally an infinite number of examples. If x is discrete then the probability of finding x is equal to the fractions of the total number of examples in which x is found. If x is a continuous variable then the probability of finding x in dx centered at x_o

$$P(x_o)\,dx$$

is defined as the fraction of the total number of examples in which x is found in the desired interval.

Entropy and Probability

Let us consider a large number of systems (called an ensemble of systems) which are isolated from each other and which are identical as to the number of particles they contain, and as to their total energy. They might also be identical as to volume, shape of the volume, linear momentum, angular momentum and perhaps to some other unspecified quantities. Let x be some quantity associated with the systems which is not fixed and can vary from system to system. An example is the density of particles in some little region of the container. It should be stressed that x can be almost any quantity that can be associated with the system. We can even consider sets of x's which contain many terms (for example the densities in many subregions of the system). However the number o x's should be much less than the number of particles in the system. The x's should not give a complete dynamical description of the system. Let us further assume that the ensemble is an equilibrium ensemble by which we mean that the number of systems found in any given possible state at one time is the same as that found at any other time. Then it is a basic result of equilibrium statistical mechanics that the probability of finding a system with x lying between x and $x + dx$ is given by

$$P(x)\,dx \propto e^{S(x)/k}\,dx$$

or

$$S = k\ell nP \;,$$

where S is the entropy of the system when x takes on the prescribed value. If x_o is the equilibrium value of x, or the most probable value of x, then for small derivations from this value we have

$$\xi = x - x_o$$

$$P(\xi)\, d\xi \propto e^{S(x_o)/k}\, e^{\frac{1}{k}\frac{\partial^2 S}{\partial x^2}\xi^2/2} \propto e^{\frac{1}{k}\frac{\partial^2 S}{\partial x^2}\xi^2/2}$$

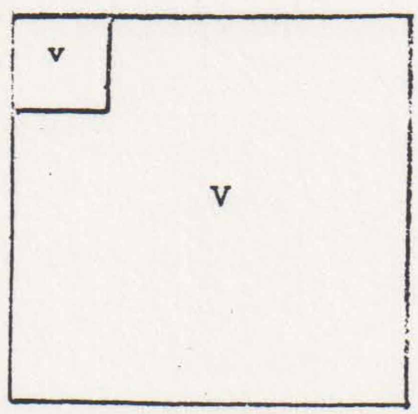
$$\left(\frac{\partial S}{\partial x}\right)_{x = x_o} = 0 \; .$$

Since S is a maximum for $x = x_o$, $\frac{\partial^2 S}{\partial x^2} < 0$. If we are considering more than one x then we would have

$$\xi_1 = x_1 - x_{10}, \xi_2 = x_2 - x_{20}, \ldots, \xi_m = x_m - x_{mo}$$

$$P(\xi_1, \ldots, \xi_m) \propto e^{\frac{1}{k}\sum_{\ell m}\frac{\partial^2 S}{\partial x_\ell \partial x_m}\frac{\xi_\ell \xi_m}{2}} \; .$$

Problem: Consider a system of volume V containing N particles. Consider a little volume v contained inside.



Use the formula that the entropy density for an ideal gas is

$$s = - nk\, \ell n\, n$$

to find the probability that there are n_1 particles in volume v and $N-n_1$ in volume V - v .

Fluctuations and the Potential of the Average Force

The entropy is a state variable, which means that if we know the ~~entropy~~ energy, the volume, the total number of particles contained, and the values of all the x's, we would know the entropy. Thus

$$S = S(u, V, N, x_1, \ldots, x_m) \; . \tag{1}$$

Thus we have the differential relation

$$dS = \left(\frac{\partial S}{\partial u}\right) du + \left(\frac{\partial S}{\partial V}\right) dV + \left(\frac{\partial S}{\partial N}\right) dN + \sum_m \frac{\partial S}{\partial x_m}\, dx_m \; . \tag{2}$$

Now from thermodynamics we have

$$\frac{1}{T} = \frac{\partial S}{\partial u} \tag{3}$$

$$+P = \left(\frac{\partial S}{\partial V}\right) \Big/ \partial S/\partial u \tag{4}$$

$$\mu = \left(\frac{\partial S}{\partial N}\right) \Big/ \partial S/\partial u \; , \tag{5}$$

so

$$TdS = du + PdV - \mu dN + \sum_m X_m\, dx_m \; , \tag{6}$$

where

$$X_m = + \frac{\partial S}{\partial x_m} \Big/ \partial S/\partial u = X_m(u, V, N, x_1, \ldots, x_m) \; . \tag{7}$$

The entropy associated with a fluctuation of x_i from its most probable value is thus given by

$$\Delta S = \int_{S_o}^{S} dS = + \sum_i \int_{x_{mp}}^{x_i} \frac{X_i\, dx_i}{T} \; . \tag{8}$$

There is no change in u, V or N since these quantities are fixed for the system. Now to gain a physical understanding of this relation let us imagine that instead of allowing the fluctuation to arise from the random motion of the particles we stepped in and forced the system into this state. We further imagine that we do it very slowly so that the motion is adiabatic and hence we do not increase the entropy of the system. Thus in doing this dS is zero but the energy of the system is no longer constant since we are doing work on it from outside. We then have

$$du + \sum_i X_i \, dx_i = 0 \tag{9}$$

or

$$du = - \sum_i X_i \, dx_i \quad . \tag{10}$$

Thus $X_i \, dx_i$ is the change in energy of the system; X_i may be thought of as a force associated with x_i . We then have

$$\delta u = \Phi = \sum_i \int X_i \, dx_i \; . \tag{11}$$

Now if the change in energy associated with the x_i is very small compared to the total energy of the system then any changes in temperature associated with it will be small and T can be considered constant. Then we may relate (11) to the change in entropy given by Eq. (8)

$$\Delta S = - \frac{\Phi}{T} \; .$$

The probability of finding the state $x_1, \ldots, x_m$ is then given by

$$P(x_1, \ldots, x_m) \propto e^{-\frac{\Phi(x_1, \ldots, x_m)}{kT}} \, dx_1 \, dx_2, \ldots, dx_m$$

where Φ is the work required to create the fluctuation by means of external forces.

<u>Problem:</u> Consider a gas of neutral particles in a cubic box of volume $L^3 = V$. Let n_k be the amplitude of the k'th Fourier mode of the density

$$n_k = \int n(\underset{\sim}{r}) \, e^{ik \cdot r} dr \qquad \begin{aligned} k_x &= \frac{2\pi \ell}{L} & \ell &= 1, 2, \ldots \\ k_y &= \frac{2\pi n}{L} & n &= 1, 2, \ldots \\ k_z &= \frac{2\pi m}{L} & m &= 1, 2, \ldots \end{aligned} \quad .$$

Find the distribution function for n_k, $P(n_k)$. Find $P(n_k, n_{k'})$ $k' \neq k$. Are n_k and $n_{k'}$ correlated?

Consider a plasma consisting of electrons and singly charged ions contained in a cubic box of volume $L^3 = V$. Find the distribution function for the k'th Fourier components of the density of electrons and ions. Are they independent? Suppose there is also a neutral component to the plasma. What is the distribution function for the k'th Fourier component of its density? Is it independent of the fluctuations in electron and ion density? What is the distribution function for this k'th component of the electric field?

IRREVERSIBLE STATISTICAL MECHANICS

REFERENCES

"Selected Papers on Noise and Stochastic Processes", edited by Nelson Wax, Dover 1954.

"Statistical Physics", Landau-Lifshitz, Addison-Wesley Publishing Co., Inc., 1958.

"Brownian Motion", A. Einstein, Dover.

"Thermodynamics of Irreversible Processes", S. R. DeGroot, Interscience, New York 1951.

"The Principles of Statistical Mechanics", R. C. Tolman, Clarendon Press, Oxford 1938.

"Mathematical Theory of Nonuniform Gases", S. Chapman and T. S. Cowling, Cambridge 1939, also Dover.

"Kinetic Theory of Gases", Jeans, Dover.

"Elementary Statistical Physics", Kettel, John Wiley and Sons, Inc. 1958, 1961.

Mathematical Review

Fourier Transform

The Fourier transform $g(k)$ of a function $f(x)$ is defined by

$$g(k) = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \tag{1}$$

If $f(x)$ is a real function then

$$g(-k) = g^{*}(k) \tag{2}$$

We may write $f(x)$ in terms of its Fourier transform

$$f(x) = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} g(k)\, e^{ikx}\, dk \tag{3}$$

Also we have the relation

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-y)}\, dk = \delta(x-y) \tag{4}$$

where $\delta(x-y)$ is the Dirac delta function.

We may write $e^{ikx} = \cos kx + i \sin kx$. If $f(x)$ is even the integral over the sin terms in (1) give 0 and we have

$$g(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \cos kx\, dx \tag{5}$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k) \cos kx\, dk \tag{6}$$

Parsival's Theorem

Parsival's theorem states

$$\int_{-\infty}^{\infty} |g(k)|^2\, dk = \int_{-\infty}^{\infty} |f(x)|^2\, dx \tag{7}$$

Proof -

$$|g(k)|^2 = g(k)\, g^*(k) \tag{8}$$

$$\int_{-\infty}^{\infty} g(k)\, g^*(k)\, dk = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} f(\xi)\, e^{-ik\xi}\, d\xi \int_{-\infty}^{\infty} f^*(x)\, e^{ikx}\, dx$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi dx\, \delta(x - \xi)\, f(\xi)\, f^*(x) = \int_{-\infty}^{\infty} f(x)\, f^*(x)\, dx \tag{9}$$

Example -

Let $f(x) = e^{-\alpha x^2}$

$$g(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx - \alpha x^2}\, dx \tag{10}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{- \alpha \left(x + \frac{ik}{2\alpha}\right)^2 - \frac{k^2}{4\alpha}\right\} dx \tag{11}$$

$$= \frac{1}{\sqrt{2\pi}}\, e^{-\frac{k^2}{4\alpha}} \tag{12}$$

Moments

Let $P(x)\,dx$ be the probability that x lies between x and $x + dx$. The average value of x is given by

$$\bar{x} = \int x\, P(x)\, dx \tag{13}$$

Denote the deviation of x from its average value by δx

$$\delta x = x - \bar{x} \tag{14}$$

The mean square value of δx is given by

$$\overline{\delta x^2} = \int (x - \bar{x})^2 \, P(x) \, dx \tag{15}$$

$$= \int (x^2 - 2x\bar{x} + \bar{x}^2) \, P(x) \, dx = \int (x^2 - \bar{x}^2) \, P(x) \, dx$$

$$= \overline{x^2} - \bar{x}^2$$

$\overline{\delta x^2}$ is called the second moment of the distribution $P(x)$ about its mean value $\bar{x}$. $\overline{\delta x^n}$ is called the n th moment of $P(x)$ about its mean value. In general if one knows all the moments of a function, one can construct the distribution.

To prove this, take the Fourier transform of $P(x)$

$$u(k) = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} P(x) \, e^{-ikx} \, dx \tag{16}$$

Now take the n th derivative of $u(k)$ with respect to k

$$\frac{d^n u(k)}{d k^n} = \frac{(-i)^n}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} x^n \, P(x) \, e^{-ikx} \, dx \tag{17}$$

If we evaluate this expression for $k = 0$, we get

$$\left. \frac{d^n u(k)}{d k^n} \right|_{k=0} = \frac{(-i)^n}{(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} x^n \, P(x) \, dx = \frac{(-i)^n \, \overline{x^n}}{(2\pi)^{\frac{1}{2}}} \tag{18}$$

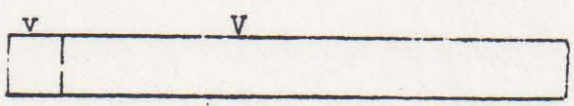
Thus if we know all $\overline{x^n}$ we know all derivatives of $u(k)$ at $k = 0$. We can construct $u(k)$ from its taylor series and the inverse transform gives $P(x)$.

$$u(k) = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{d^n u(k)}{d k^n} \right|_{k=0} k^n \qquad (19)$$

While in principle it is possible to do this, in practice it is not always useful since we need to know all the moments which we usually don't. Further, all derivatives must exist and $u(k)$ must be analytic, i.e., expandable in a power series.

Stochastic or Random Variable

A stochastic or random variable is a variable which can take on a definite set of values and it attains each one with a definite probability. For an example consider the following system



It consists of a small volumeconnected to a resevoir through a small orifice. Let the total number of particles be N. Then the number of particles in the small volume is a random variable and can take on any integer value from 0 to N. If the probability of finding a particle in a small volume dv is proportional to dv and independent of where the volume element is located, then the probability of finding n particles in v is

$$P(n) = \frac{v^n V^{N-n}}{V_T^N} \frac{N!}{n!(N-n)!} \qquad (20)$$

It is also possible for a random variable to take on a continuous set of values. An example of such a variable is the velocity of an air molecule. In this case we can not give the probability that the particle has a definite velocity, but rather must give the probability that the velocity lies in some velocity interval, say between v and v + dv. For a Maxwell velocity distribution this probability is given by

$$P\,(v)\,dv = \frac{1}{\sqrt{2\pi}\;v_o}\;e^{-\frac{v^2}{2\,v_o^{\,2}}}\;dv \qquad (21)$$

Random Process

A random process is a process which can not be completely determined by knowledge available to us. It is a process which if repeated or duplicated to the best of our ability, will not give exactly reproducible results. All that we can know about random processes are the probabilities that certain events will occur.

Suppose a certain variable $n(t)$ characterizes a random process. For example, it might be the number of atoms in the small volume just considered. The random process is the entering and leaving of the molecules. We can not predict exactly how many atoms will be in the small chamber at any given time. We can, however, give the probability that we will find n atoms at time t. Also we may give such things as the probability that given that there are n_1 atoms in the volume at time t, we will find n_2 atoms there at $t + \tau$. A random process is completely described by the following set of probability distributions:

$W_1(nt)\,dn$ probability that we find n between n and n + dn at times t (22)

$W_2(n_1t_1; n_2t_2)\,dn_1\,dn_2$ joint probability that we find n between n_1 and $n_1 + dn_1$ at time t_1 and between n_2 and $n_2 + dn_2$ at time t_2 (23)

$W_m(n_1t_1; n_2t_2; \ldots; n_mt_m)\,dn_1\,dn_2 \ldots dn_m$ (24)

joint probability that we find n between n_1 and $n_1 + dn_1$ at t_1, between n_2 and $n_2 + dn_2$ at t_2, ... between n_m and $n_m + dn_m$ at t_m.

Probability as it is used here should be understood to have the following meaning. If the random process is repeated many times (strictly speaking we should pass to the limit of an infinite number of repetitions) and an event is observed to occur for a certain fraction, (f), of these trials, then f is the probability of occurence of the event.

The W's satisfy the following simple relations:

$$(1)\ W_m \geq 0 \qquad (25)$$

Since W is a probability

$$(2)\ W_m(n_1t_1; \ldots; n_jt_j; \ldots; n_\ell t_\ell; \ldots; n_mt_m) \qquad (26)$$

$$= W_m(n_1t_1; \ldots; n_\ell t_\ell; \ldots; n_jt_j; \ldots; n_mt_m)$$

(j and ℓ interchanged). Both refer to the same set of events.

$$(3)\ \int W_m(n_1t_1; \ldots; n_mt_m)\, dn_{k+1} \ldots dn_m = \qquad (27)$$

$$W_k(n_1t_1; \ldots; n_kt_k)$$

$$k < m$$

Since each W_m must imply all previous W_k $k < m$.

<u>Types of Random Processes</u>

There are two types of random processes which we will consider, the completely random process and the Markoff process. A completely random process is one where there is no correlation between events at one time and those at another. For such a process

$$W_2(n_1t_1; n_2t_2) = W_1(n_1t_1)\, W_1(n_2t_2)\ . \tag{28}$$

The joint probability of finding n between n_1 and $n_1 + dn_1$ at t_1 and between n_2 and $n_2 + dn_2$ at t_2 is simply the product of the probability of finding n in the desired intervals at t_1 and t_2. Likewise all higher W's are simply products of W_1's. All information about the process is contained in W_1.

It is hard to think of a continuous process which is completely random. However, for processes which occur in discrete steps it is easy to find examples. For example, the flipping of a coin. The probability that heads or tails turn up on any one flip is independent of what has turned up before (provided the coin and flipper are honest). Even if heads had turned up 100 times in a row, the chances of getting a head on the 101st flip would be $\frac{1}{2}$. (Of course, if this happened, you would suspect that the probability of getting a head was not $\frac{1}{2}$. This can only be determined by flipping the coin a very large number of times [in principle an ∞ number of times] and recording the fraction of times heads show up.) The second type of random process is the so-called Markoff process. For this type of process what happens in the next instant of time depends only on the present state of the system and not on its previous history. All the information for this type of process is contained in W_2.

To make the definition of a Markoff process precise, we define the conditional probability

$$P_2(n_1t_1 \mid n_2t_2)\, dn_2 \tag{29}$$

as the probability that n lies between n_2 and $n_2 + dn_2$ at time t_2 given that it had the value n_1 at t_1.

Likewise the conditional probability

$$P_m(n_1t_1;\; n_2t_2;\; \ldots;\; n_{m-1}t_{m-1} \mid n_mt_m)\, dn_m \tag{30}$$

is the probability that we find n between n_m and $n_m + dn_m$ at time t_m given that n took on the value n_1 at t_1, n_2 at t_2, ..., n_{m-1} at t_{m-1}. We will take the t's to be such that

$$t_1 < t_2 < t_3 < \cdots < t_{m-1} < t_m. \tag{31}$$

In terms of P_2 we have for W_2

$$W_2(n_1t_1;\, n_2t_2) = W_1(n_1t_1)\, P_2(n_1t_1 \mid n_2t_2)\,. \tag{32}$$

Also P_2 has the following properties

$$P_2(n_1t_1 \mid n_2t_2) \geq 0, \tag{33}$$

and

$$\int P_2(n_1t_1 \mid n_2t_2)\, dn_2 = 1 \tag{34}$$

$$W_1(n_2t_2) = \int dn_1\, W_1(n_1t_1)\, P_2(n_1t_1 \mid n_2t_2) \tag{35}$$

A Markoff process is one where the conditional probability

$$P_m(n_1t_1;\; \ldots;\; n_{m-1}t_{m-1} \mid n_mt_m)\, dn_m \tag{36}$$

is equal to

$$P_2(n_{m-1}t_{m-1} \mid n_mt_m)\, dn_m \tag{37}$$

Thus the conditional probability that n takes on a value between n_m and $n_m + dn_m$ at t_m, given that it took on values n_1, n_2, ..., n_{m-1} at

$t_1, t_2, \ldots, t_{m-1}$ only depends on the fact that it took on the value n_{m-1} at t_{m-1} and does not depend on any of the previous values it took on.

All higher W's can be built up from W_2 for a Markoff process. For example,

$$W_3(n_1t_1; n_2t_2; n_3t_3) = W_2(n_1t_1; n_2t_2) \cdot P_2(n_2t_2 \mid n_3t_3)$$

$$= \frac{W_2(n_1t_1; n_2t_2)\, W_1(n_2t_2)\, P_2(n_2t_2 \mid n_3t_3)}{W_1(n_2t_2)}$$

$$= \frac{W_2(n_1t_1; n_2t_2)\, W_2(n_2t_2; n_3t_3)}{W_1(n_2t_2)} \qquad (38)$$

Finally, for such a process P_2 has the following property

$$P_2(n_1t_1 \mid n_2t_2) = \int P_2(n_1t_1 \mid nt)\, P_2(nt \mid n_2t_2)\, dn \qquad (39)$$

for all t such that

$$t_1 \leq t \leq t_2 .$$

Stationary Processes

In many cases we deal with what are known as stationary processes. For such a process $W_1(nt)$ is independent of t. The probability that we find a value n does not depend on when we look at the system. Likewise, the probabilities

$$W_2(n_1t_1; n_2t_2) \qquad (40)$$

$$P_2(n_1t_1 \mid n_2t_2) \qquad (41)$$

can depend only on the time difference

$$t_2 - t_1$$

since W_1 can be obtained from W_2 by integration and P_2 can not depend on when the first observation was made.

For a stationary Markoff Process we will write

$$P_2(n_1 t_1 | n_2 t_2) = P_2(n_1 0 | n_2 t) = P_2(n_1 | n_2 t) \qquad (42)$$

$$t = t_2 - t_1$$

The Correlation Function

The auto correlation function for a random variable x is defined by

$$C(\tau) = \langle x(t)\, x(t+\tau) \rangle \qquad (43)$$

where $\langle \quad \rangle$ denotes a time average over t. We may also consider another average

$$\tilde{C}(t, \tau) = \overline{x(t)\, x(t+\tau)} \qquad (44)$$

where the $\overline{\qquad}$ denotes an ensemble average. Ensemble average has the following meaning. The quantity x is a variable which describes the state of the system. If we investigate a great many systems which are as nearly alike as we can make them, x will be a definite function of time for one system, but it will be a different function for each system. The ensemble average of x $\left[\text{or of } x(t)\, x(t+\tau)\right]$ is the average value of x $\left[\text{or of } x(t)\, x(t+\tau)\right]$ over a collection or ensemble of systems. Strictly speaking, we will consider the ensemble to contain an ∞ number of systems.

A stationary random process is now one in which an ensemble of systems performing the process will appear stationary in the sense that we will always find the same number of systems in any state no matter when we look at the ensemble. Individual members of the ensemble may show a time dependence. For a stationary ensemble we have

$$\tilde{C}\,(t,\tau) = \left\langle \tilde{C}(t,\tau) \right\rangle = \overline{C(\tau)} = C(\tau)\ . \tag{45}$$

Here we assume that the correlation function is independent of the member of the ensemble, for if it weren't, the system would be undergoing different random processes.

We may also form correlation functions between two different variables, say for x and y. In this case we define the correlation function by

$$C\,(x\,y,\tau) = \left\langle x(t)\,y(t+\tau) \right\rangle \quad .$$

One can also form ensemble averages and thus define

$$\tilde{C}\,(xy,t,\tau) = \overline{x(t)\,y(t+\tau)}\ .$$

For a stationary ensemble $\tilde{C}$ is independent of t and

$$\tilde{C}\,(xy,t,\tau) = C\,(xy,\tau)\ .$$

The correlation function $C\,(xy,\tau)$ need not equal $C(yx,\tau)$ even for stationary processes.

One can also form correlation function with respect to spatial positions rather than time. Thus for the case of auto correlation we have

$$C\,(\underset{\sim}{\rho}) = \left\langle x(\underset{\sim}{r})\,x(\underset{\sim}{r}+\underset{\sim}{\rho}) \right\rangle_{\text{averaged over space}}$$

$$\tilde{C}\,(\underset{\sim}{r},\underset{\sim}{\rho}) = \overline{x(\underset{\sim}{r})\,x(\underset{\sim}{r}+\underset{\sim}{\rho})} \qquad \text{ensemble average}\ .$$

If the systems are infinite and uniform then $\tilde{C}\,(\underset{\sim}{r},\underset{\sim}{\rho})$ is independent of $\underset{\sim}{r}$ and is equal to $C(\underset{\sim}{\rho})$.

We may also correlate different quantities with respect to position just as we did with respect to time. Further we may combine time and space correlations.

Example - Shot Noise

Let us consider the problem of shot noise. Some examples of it are the following processes: the number of electrons emitted by the filament of a vacuum tube per unit time; the number of rain drops striking a unit area in a unit time in a steady rain; the number of molecules striking a small area of a container wall in a unit time; and the number of molecules coming out of a small orifice in the wall of a vessel.

We assume that the number of particles emitted (to take the emission of electrons for example) during any one time to be independent of the number emitted at any other time. Thus the emission of electrons is a completely random process.

The average number of electrons emitted per second is

$$< n > = \lim_{T\to\infty} \frac{N}{T} \qquad \text{time average} \qquad (46)$$

N is the total number of electrons emitted.

or

$$\bar{n} = \lim_{M\to\infty} \frac{\sum_{i=1}^{M} n_i}{M} \qquad \text{ensemble average} \qquad (47)$$

n_i is the number emitted in a unit of time by the ith system.

Let us ask what is the probability that K electrons will be emitted during a time t. Let us take the ensemble point of view. Let there be M systems. Let M (o, t) be the number of systems which have not emitted any particles in time o to t. M (1, t) is the number of systems which have emitted 1 electron in the time 0 to t. M (K, t) is the number of systems which have emitted K in this time. Now the change in M (o, t) in dt is equal to $-\bar{M}(o, t)$ times the probability that each system will emit 1 particle during dt

$$\frac{d\,M(o,t)}{dt} = -\bar{n}\,M(o,t)\,. \tag{48}$$

Likewise

$$\frac{d\,M(K,t)}{dt} = -\bar{n}\,M(K,t) + \bar{n}\,M(K-1,t)\,. \tag{49}$$

We can write this equation

$$\frac{d\,M(K,t)}{dt} + \bar{n}\,M(K,t) = \bar{n}\,M(K-1,t)\,, \tag{50}$$

$$\frac{d}{dt}\,e^{\bar{n}t}\,M(K,t) = \bar{n}\,e^{\bar{n}t}\,M(K-1,t)\,, \tag{51}$$

$$M(K,t) = \int_0^t \bar{n}\,e^{\bar{n}(\tau-t)}\,M(K-1,\tau)\,d\tau\,. \tag{52}$$

First, $M(o,t) = M\,e^{-\bar{n}t}$.

Substituting in the above equations gives

$$M(1,t) = M\int_0^t \bar{n}\,n^{-\bar{n}t}\,d\tau = M\,t\,\bar{n}\,e^{-\bar{n}t} \tag{53}$$

$$M(2,t) = \int_0^t n^2\quad e^{-\bar{n}t}\,d\tau = M\,\frac{(\bar{n}\,t)^2}{2}\,e^{-\bar{n}t} \tag{54}$$

$$M(K,t) = M\,\frac{(\bar{n}\,t)^K\,e^{-\bar{n}t}}{K!} \tag{55}$$

$$P(K, t) = \frac{M(K, t)}{M} = \frac{(\bar{n}\, t)^K e^{-\bar{n}t}}{K!} \quad . \tag{56}$$

What is the most probable number emitted in a time t. Take the $\ln$ of $P(K, t)$

$$\ln P(K, t) \cong K \ln \bar{n}\, t - K(\ln K - 1) \tag{57}$$

$$\frac{d \ln P(K, t)}{dK} = \ln \bar{n}\, t - \ln K + 1 - 1 \tag{58}$$

$$\underset{\substack{\text{most}\\ \text{probable}}}{K} = \bar{n}\, t \quad ,$$

Let us find the probability that a number different from $K_{\substack{\text{most}\\ \text{probable}}}$ is emitted. Let

$$\Delta K = K - \bar{n}\, t$$

$$\frac{d^2 \ln P(K, t)}{d K^2} = -\frac{1}{K} \tag{59}$$

$$\ln P(\Delta K, t) = \ln P(K_{MP}, t) - \frac{\Delta K^2}{2\bar{n}t}$$

$$P(\Delta K, t) \; \alpha \; e^{-\frac{\Delta K^2}{2\bar{n}t}} \quad . \tag{60}$$

This is an example of the central limit theorem.

The average number of electrons emitted during the time o to t is given by

$$\overline{K(t)} = \sum_{K=0}^{\infty} \frac{K\, e^{-\bar{n}t} (\bar{n}\, t)^K}{K!} = e^{-\bar{n}t} \sum_{K=0}^{\infty} \frac{K\, (\bar{n}\, t)^K}{K!}$$

$$= \bar{n}\, t\, e^{-\bar{n}t} \sum_{K=1}^{\infty} \frac{(\bar{n}\, t)^{K-1}}{(k-1)!} = \bar{n}\, t\, e^{-\bar{n}t}\, e^{\bar{n}t} = \bar{n}\, t \; . \qquad (61)$$

The mean and most probable value are the same.

Exercise

Compute the deviation from Gaussian for $P(K, t)$ by carrying the expansion one step higher, How much error is made if $\bar{n}\, t = 10, 100, 1000$ and $\Delta K^2 = \bar{n}\, t,\ 10\, \bar{n}\, t,\ 100\, \bar{n}\, t$.

If K are emitted in the time o to t, what is the probability that K′ are emitted between O and $t + \tau$. Since the process is a completely random process, we have

$$P_2 (K\, t \mid K'\, t+\tau) = \frac{e^{-\bar{n}t} (\bar{n}\, \tau)^{K'-K}}{(K'-K)!} \qquad K' \geq K \qquad (62)$$

$$0 \qquad K' < K$$

$$W (K\, t;\, K'\, t+\tau) = \frac{e^{-\bar{n}\,(t+\tau)} (\bar{n}\, t)^K (\bar{n}\, t)^{K'-K}}{K!\, (K'-K)\, !} \qquad (63)$$

$K' \geq K$.

Let us compute the average number of electrons emitted in the time o to $t+\tau$ if K are emitted in the time 0 to t.

$$\tilde{K} = \sum_{K'=K}^{\infty} \frac{K'\, e^{-\bar{n}\tau} (\bar{n}\, \tau)^{K'-K}}{(K'-K)!} \; . \qquad (64)$$