CONTENTS

CHAPTER 1

INTRODUCTION

Reasons for Studying Fluctuations in Semi-conductors

The expression 'fluctuation phenomena in semi-conductors' refers to the spontaneous fluctuations in the current passing through, or the voltage developed across, semi-conductor samples or devices. A more common name is 'noise in semi-conductors', which refers to the acoustical noise heard when these electrical fluctuations are amplified by an audio amplifier and fed into a loudspeaker. At present this name is generally applied to any fluctuating current and/or voltage, even if no acoustical effect is involved.

A study of the fluctuation phenomena in semi-conductors and semi-conductor devices is important for the following reasons:

(1) These fluctuations reflect in part the atomistic character of the conduction mechanism and, as such, their study is basic for a thorough understanding of the behaviour of these materials and devices.

(2) Once these fluctuation phenomena are understood, they may be used as a new tool in the study of the physics of these materials and devices. In particular, it may make a certain physical phenomenon more clearly detectable, or it may allow a physical quantity of the material or the device to be determined with greater accuracy than that obtainable by other methods.

(3) Semi-conductor materials and/or devices are mainly used to measure small physical quantities or to amplify small signals. The spontaneous fluctuations in current and voltage set a lower limit to the quantities which may be measured or to the signals which may be amplified. It is important to know the factors contributing to these limits and to apply this knowledge in order to find the optimum operating conditions of these materials or devices. Finally, it is important to find out how the materials or the devices can be improved to lower this limit.

It is presupposed that the reader is familiar with current semi-conductor nomenclature, the diffusion equations and the concepts of donors, acceptors, traps and recombination centres.

Classification of Various Sources of Noise

In classifying the noises observed in semi-conductor materials and semi-conductor devices, distinction is usually made between thermal noise, flicker noise and shot noise. Thermal noise occurs in any conductor and is caused by the random motion of its current carriers. Flicker noise is distinguished by its peculiar spectral intensity which is of the form const.$/f^{\alpha}$ with α close to unity; in this respect it resembles flicker noise in vacuum tubes, hence the name. Shot noise has the main characteristic of the spectrum, being white at the lower frequencies, in which respect it resembles shot noise in vacuum tubes, hence the name. In the earlier literature, no distinction was made between shot noise and flicker noise, so that theories developed for the former also applied to the latter. Moreover, these names were introduced on a more or less heuristic basis, without sufficient reference to the physical causes of the noise, and for this reason the following terminology is suggested.

(*a*) *Generation–recombination noise*, which is caused by spontaneous fluctuations in the generation rates, recombination rates, trapping rates, etc., of the carriers, thus causing fluctuations in the free carrier densities. In bulk material the name generation–recombination noise is more appropriate than shot noise, since these carrier density fluctuations exist even if no electric field is applied; applying a field is only the most convenient way of detecting the fluctuations. Junction devices, such as diodes and transistors, operate along the principle of minority carrier injection and the applied voltage is needed to produce or change the injection level. The generation–recombination noise in these devices therefore shows much closer resemblance to shot noise than in the bulk material case (see Chapters 8 and 9).

(*b*) *Diffusion noise*, caused by the fact that diffusion is a random process; consequently fluctuations in the diffusion rate give rise to localized fluctuations in the carrier density. In bulk material it is the cause of thermal noise: in junction devices it gives a major contribution to shot noise.

(*c*) *Modulation noise*, which refers to noise not directly caused by fluctuations in the transition or diffusion rates but which, instead, is due to carrier density fluctuations or current fluctuations caused by some modulation mechanism (examples are discussed in Chapters 5 and 6).

CHAPTER 2

CHARACTERIZATION OF NOISINESS IN TWO- AND FOUR-TERMINAL NETWORKS

NOISE CHARACTERIZATION IN TWO-TERMINAL NETWORKS [1]

IN a two-terminal network the noise in a frequency interval Δf can always be characterized by an equivalent e.m.f. $\sqrt{\overline{e^2}}$ in series with the device or by an equivalent current generator $\sqrt{\overline{i^2}}$ in parallel to the device; $\sqrt{\overline{e^2}}$ represents the open-circuit noise e.m.f. and $\sqrt{\overline{i^2}}$ the short-circuit noise current of the network, both quantities usually being expressed per unit bandwidth. It is often convenient to use other quantities to characterize the noise.

The equivalent noise resistance R_n is defined by:

$$\overline{e^2} = 4kTR_n\Delta f \qquad \ldots.(2.1)$$

where T is room temperature and k is Boltzmann's constant. According to Nyquist's theorem, the network itself and a resistance R_n at the temperature T have the same mean square value of the open-circuit noise e.m.f. for the bandwidth Δf.

The equivalent saturated diode current I_{eq} is defined as.

$$\overline{i^2} = 2eI_{eq}\Delta f \qquad \ldots.(2.2)$$

According to Schottky's theorem, the network and a saturated diode carrying a current I_{eq} (and showing full shot noise only) have the same mean square value of the short-circuit noise current for the bandwidth Δf.

If $Z = R + jX$ is the impedance of the device and $Y = 1/Z = g + jb$ its admittance, then the noise ratio n of the device is defined as:

$$n = \frac{\overline{e^2}}{4kTR\Delta f} = \frac{\overline{i^2}}{4kTg\Delta f} \qquad \ldots(2.3)$$

Besides the noise ratio n the equivalent noise temperature T_n may also be introduced by the definition:

$$T_n = nT \qquad \ldots.(2.3a)$$

The importance of the quantities n and T_n is that they allow the available noise power P_{av} for the bandwidth Δf to be expressed as:

$$P_{av} = \frac{\overline{e^2}}{4R} = \frac{\overline{i^2}}{4g} = nkT\Delta f = kT_n\Delta f \qquad \ldots (2.4)$$

The available power is thus equal to n times the available noise power of a resistance at room temperature T.

The following relations are easily proved:

$$n = \frac{R_n}{R} = \left(\frac{e}{2kT}\right)\frac{I_{eq}}{g} \simeq 20\frac{I_{eq}}{g}; \; I_{eq} = \left(\frac{2kT}{e}\right)R_n|Y|^2 \simeq \frac{1}{20}R_n|Y|^2 \qquad \ldots (2.5)$$

Knowing the quantities R_n, I_{eq} and n is often an important step in determining the exact cause of the noise.

Noise Characterization in Four-terminal Networks[1-3]

In an active four-terminal network, two noise sources are needed to characterize the noise. In an equivalent π representation the most general circuit has three noise current generators I_{ab}, I_{ac} and I_{bc} connected between the three nodal points a, b and c. But this

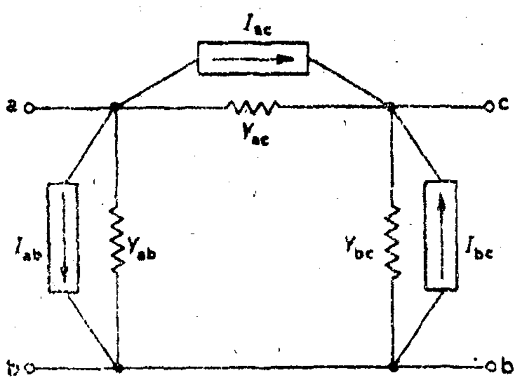

Figure 2.1. Equivalent active π network with the noise characterized by three current generators I_{ab}, I_{ac} and I_{bc}

can be transformed into an equivalent circuit consisting of a current generator $(I_{ab} + I_{ac})$ connected between a and b and a current generator $(I_{ac} + I_{bc})$ connected between b and c (see *Figure 2.1*). In an equivalent T representation the most general circuit consists of three noise e.m.f.'s in series with the three leads; this can be reduced to an equivalent circuit with only two noise e.m.f.'s in a similar manner. This gives the following possibilities:

(1) Two current generators i_1 and i_2 connected across input and output, respectively.

(2) Two noise e.m.f.'s e_1 and e_2 connected in series with input and output, respectively.

(3) One noise e.m.f. e and one current generator i: this case has four possible configurations, the most important ones being:

(*a*) noise e.m.f. in series with input: noise current generator is parallel to output;

(*b*) noise e.m.f. in series with input: noise current generator is parallel to input.

Generally, the two noise sources are partially correlated (see Chapter 3). For a proper determination of the noise behaviour, four network parameters must be determined; in case 1, for example, $\overline{i_1^2}$, $\overline{i_2^2}$ and the real and imaginary parts of $\overline{i_1^* i_2}$ must be known. In some simple cases, $\overline{i_1^* i_2} = 0$ (the noise sources are then independent), thus reducing the four quantities to two; this occurs in transistors at relatively low frequencies (see Chapter 9). It may also happen that one of the quantities is zero, thus reducing the four quantities to only one, which may be characterized by an equivalent noise resistance; this occurs in vacuum tubes at medium frequencies.

The noisiness of an active four-terminal network can now be characterized in several ways. Let the signal source have an internal impedance $Z_s = (R_s + jX_s)$ or an internal admittance $Y_s = 1/Z_s = (g_s + jb_s)$. For example, the noise may be represented by an equivalent e.m.f. e_n in series with the source or by an equivalent noise current generator i_n in parallel to the source; these quantities are defined so that the output noise power of the network is doubled if the noise e.m.f. e_n or the noise current generator i_n are introduced. One may then define the equivalent noise resistance R_n or the input equivalent saturated diode current I_n of the network by the equations:

$$\overline{e_n^2} = 4kTR_n\Delta f; \quad \overline{i_n^2} = 2eI_n\Delta f \qquad \ldots (2.6)$$

where T is room temperature, k is Boltzmann's constant, e is the electron charge, and Δf a small frequency interval. Both quantities R_n and I_n may depend upon the internal impedance of the source.

It is more common to introduce the noise figure F of the network as the ratio of the total output noise power over the output noise power due to the thermal noise of the source. The latter can be represented by a noise e.m.f. $\sqrt{4kTR_s\Delta f}$ in series with the source or by a noise current generator $\sqrt{4kTg_s\Delta f}$ in parallel to the source.

Then, according to equation (2.6):

$$F = \frac{R_n}{R_s} = \frac{e}{2kT}\frac{I_n}{g_s} \quad \ldots (2.7)$$

The noise figure always shows a parabolic dependence on R_s and has a minimum value F_{min} for $R_s = (R_s)_{min}$.

The smallest available signal power which may be detected against the noise background of an amplifier of noise figure F and bandwidth B is about $FkTB$. The noise figure F should therefore be made as small as possible under the existing operating conditions.

In many cases it is possible to change the source impedance, as viewed from the input of the amplifier, within a wide range with the help of a lossless matching network. In such cases the amplifier with the lowest value of F_{min} is the best one. In other cases it is necessary to connect the signal source direct to the amplifier without the benefit of a lossless matching network, in which case the amplifier with the lowest value of F_{min} may be a rather poor choice. In the case of a low-impedance signal source the amplifier with the lowest noise resistance R_n is the best one, whereas with a high-impedance signal source the amplifier with the lowest input equivalent saturated diode current I_n is preferred[4].

The first step in characterizing the noisiness of an amplifier stage is to find the noise sources of its active element and locate the proper positions of these sources in its equivalent circuit. It is then possible to determine the most suitable operating conditions of a given active element, or to design the active element so that it gives the lowest noise figure F under the existing operating conditions (see Chapter 11).

The representation of the noise properties of an active network by an equivalent circuit is not unique, since a given circuit can be transformed into another one by applying certain network theorems. Usually that equivalent circuit fitting closest to the physics of the device is sought (see Chapter 9).

NOISE-MEASURING EQUIPMENT[1]

Small noise signals are amplified and detected, thus enabling them to be measured. The first concern of a noise investigation, therefore, is to build a linear low-noise amplifier which will amplify the noise signals in the desired frequency bands to a level high enough to be detected.

An important requirement is that the amplifier be linear: to illustrate this, first consider the case where the noise signal is detected by a truly linear detector. The response of such a detector is proportional to the square root of the output noise power; the over-all response of amplifier plus detector only obeys a simple law if the amplifier is perfectly linear. Next, consider the case where the noise signal is detected by a truly quadratic detector. The output of the detector is now directly proportional to the output noise power and hence the output of the detector is only proportional to the input noise power if the amplifier is perfectly linear.

The most desirable detector is a quadratic detector, its great advantage being that noise signals from different sources are added quadratically and that the output of the quadratic detector is proportional to the square of the output noise signal. In order to obtain the mean square value of the wanted noise signal the contributions of the unwanted noise signals are simply subtracted from the output reading of the detector.

The best test of the quadraticity of the combination amplifier plus detector is to connect a saturated diode in parallel to the input of the amplifier and to plot the output reading of the detector against the saturated diode current I_d. If the relationship is perfectly linear, the combination amplifier plus detector is perfectly quadratic, since the mean square value of the input signal is proportional to I_d, according to Schottky's theorem*.

Several quadratic detectors are discussed in the literature: one of the most reliable ones is a combination of a thermocouple and a sensitive millivoltmeter. As long as the input to the thermocouple is sufficiently small, the combination is accurately quadratic.

A very important requirement is that the amplifier should have low noise. The most desirable situation is that the noise of the amplifier is negligible in comparison with the noise signal being detected, in which case further improvement of the amplifier is useless. If the noise of the amplifier is comparable to the noise being measured, it becomes especially important to lower the noise level of the amplifier; this can often be achieved by a careful design of the first amplifier stage and by a careful selection of the first amplifier tube or transistor.

Finally, it must be confirmed that the noise-measuring equipment can cover a sufficiently wide frequency range. Perhaps the most useful information about a noise signal is contained in the frequency dependence of its spectrum; for this reason as much of the information as possible should be sought by extending the measurements

* A noise signal contains instantaneous amplitudes far greater than its mean square value; hence, if the linearity of an amplifier is tested with a sinusoidal signal, one requirement is that the amplifier be linear for r.m.s. signal levels which are three times higher than the maximum r.m.s. value of the noise signal being detected; only then can it be certain that the amplifier responds linearly to the noise.

over a wide enough frequency range. This may be achieved either by measuring the noise at a number of fixed frequencies or by making the amplifier tunable. The latter is easily obtained by using the front end and the i.f. section of a communications receiver* as part of the equipment.

An important problem is the choice of the bandwidth B of the amplifier or receiver and the time constant τ of the detector plus indicating instrument. In the case of frequency-dependent spectra the bandwidth B should not be chosen so that the spectral intensity varies markedly over the frequency range. On the other hand, the probable error of a single reading of a noise measuring device with a bandwidth B and a detector of time constant τ is $1/\sqrt{2B\tau}$. For this reason B is often made as large as possible and the time constant τ then determined so that the probable error in a single reading is tolerable†.

Noise Measurements[1–3, 5, 6]

As outlined on pp. 3–4, noise measurements in two-terminal networks amount to the determination of the equivalent saturated diode current I_{eq}, the equivalent noise resistance R_n or the noise ratio n. In four-terminal networks the most significant quantity is the noise figure F; for a good understanding of the various factors which contribute to the noise figure, however, it may be convenient to treat the network as a two-terminal one and to determine quantities such as the equivalent output saturated diode current with the input open or short-circuited, etc.

To measure noise, a standard signal source is used in the form of a known noise source, such as a hot wire or a saturated diode. A sinusoidal signal source can also be used, but in that case a reliable attenuator is needed and the effective bandwidth of the receiver has to be determined. We discuss here the application of a saturated diode and of a sinusoidal signal source as a standard source.

* The receiver usually needs to be modified somewhat. It is necessary, for example, to disable the automatic volume control in order to obtain an output voltage proportional to the input signal. The i.f. signal can easily be brought outside by replacing the i.f. detector stage by a cathode follower stage. Finally, it often happens that the receiver is either too noisy or that its input cannot be matched to the noise source under test, in which case a low-noise pre-amplifier having the required output impedance must be provided; in some cases a good cathode follower stage using a high g_m tube is sufficient.

† The error comes about because the rectified noise signal contains low-frequency noise components which cause a fluctuating reading of the indicating instrument inversely proportional to $\sqrt{B}$; this fluctuating reading can be reduced by increasing τ as indicated.

To measure the equivalent saturated diode current, I_{eq}, of the noise source, a saturated diode is connected in parallel to it; for the sake of simplicity it is assumed that the noise of the receiver is negligible in comparison with the noise of the source under test. First, the output noise power of the amplifier, due to the source, is measured and then the saturated diode current is adjusted so that the output noise power of the amplifier is doubled. The source can be represented by a current generator $\sqrt{2eI_{eq}\Delta f}$ in parallel to the input and the diode is represented by a current generator $\sqrt{2eI_d\Delta f}$; since both sources give equal output power, $I_{eq} = I_d$. This result is independent of the bandwidth of the receiver, the only restriction being that the internal resistance, R_d, of the saturated diode is large in comparison with the source resistance R_s; this condition may not be satisfied if R_s is large*. If the source impedance, R_s, is also known, the noise ratio and the noise resistance may be determined from equation (2.5).

If a sinusoidal signal source is used, the signal must be attenuated accurately to a much lower (r.m.s.) level E_1 and a resistance R_1, large in comparison with the output resistance of the attenuator, then connected in series with it; this signal source acts as a current generator $I_1 = E_1/R_1$ in parallel to R_1. At the frequency f let the output voltage of the receiver due to the sinusoidal signal be $E_0(f)$ and let the transfer function of the receiver be defined as $g(f) = E_0(f)/I_1(f)$, where $g(f) = g_0$ at the centre frequency $f = f_0$ of the transmitted band. If the input source is tuned to the frequency f_0 and if E_1 is adjusted so that the output noise power is doubled, then

$$\frac{E_1^2}{R_1^2} g_0^2 = \int_0^\infty 2eI_{eq}\,\mathrm{d}f\,|\,g^2(f)\,| = 2eI_{eq}B_{eff}g_0^2 \qquad \ldots(2.8)$$

where B_{eff} is the effective bandwidth, defined as:

$$B_{eff} = \frac{1}{g_0^2}\int_0^\infty |\,g^2(f)\,|\,\mathrm{d}f \qquad \ldots(2.8a)$$

* R_d is caused by the Schottky effect in the diode. It may be shown that R_d is inversely proportional to I_d. Account must now be taken of the fact that the total resistance of the input circuit with the diode 'on' is less than that with the diode 'off'. Doubling the noise power gives in this case:

$$2(2eI_{eq}\Delta f)R_s^2 = (2eI_{eq}\Delta f + 2eI_d\Delta f)\left(\frac{R_sR_d}{R_s + R_d}\right)^2$$

so that:

$$I_{eq} = \frac{I_d}{1 + 4(R_s/R_d) + 2(R_s/R_d)^2} \simeq \frac{I_d}{1 + 4R_s/R_d}$$

Often R_d is so large that $I_{eq} = I_d$ in good approximation. If this is not the case, the above equation may be used to correct for the finite value of R_d.

B_{eff} may be determined either by integrating the response curve of the receiver or by comparing the sinusoidal signal with a standard noise source such as a saturated diode.

We now turn to the determination of the 'spot' noise figure F of an amplifier stage, i.e. the noise figure for a narrow-frequency band. One connects a saturated diode in parallel to the source conductance g_s at the input of the stage and connects the output of the stage to a narrow-band receiver (bandwidth less than 1/6 of that of the stage)[1]. For the sake of simplicity we assume that the noise of the receiver is negligible in comparison with the noise of the stage. If a saturated diode current I_d is needed to double the output noise power of the receiver, we have $I_d = I_n$, where I_n is the input equivalent saturated diode current. Applying equation (2.7) gives:

$$F = \frac{e}{2kT} I_d R_s \qquad \ldots .(2.9)$$

By varying R_s and measuring F as a function of R_s, one may determine the source impedance, and hence the coupling between the source and the input of the stage, for minimum noise figure.

The result thus obtained does not depend upon the bandwidth of the narrow-band receiver; if the tuning of the amplifier stage remains unaltered and the spot noise figure is measured as a function of frequency, one finds that its value depends upon frequency. Consequently the average noise figure of a wide-band receiver depends upon the bandwidth. Equation (2.9) for the spot noise figure also applies to the average noise figure, even though its derivation is no longer valid.

It has been assumed that the noise of the two-terminal source or of the four-terminal amplifier stage under test was large in comparison with the noise of the amplifier or receiver. Corrections have to be made if this is not the case, but a discussion of the problem is beyond the scope of this book[1].

References

1 VAN DER ZIEL, A., *Noise*, Englewood Cliffs: Prentice Hall, 1954

2 BECKING, A. G. T., GROENDIJK, H. and KNOL, K. S., *Philips Res. Rep.*, **10** (1955) 349

3 ROTHE, H. and DAHLKE, W., *Proc. Inst. Radio Engrs, N.Y.*, **44** (1956) 811

4 VAN DER ZIEL, A., *ibid.*, **46** (1958) 1019

5 CHAMPLIN, K. S., *ibid.*, **46** (1958) 779

6 NONNENMACHER, W., *Nachr. Tech. Z.*, **10** (1957) 559

CHAPTER 3

MATHEMATICAL METHODS

A SHORT survey is given here of the mathematical methods used in noise calculations, referring mainly to calculating averages and correlation coefficients and to making a Fourier analysis of fluctuating quantities; the results are given without proof, but for a detailed discussion, literature on the subject may be referred to[1].

PROBABILITY DISTRIBUTIONS, AVERAGES, CORRELATION

Let x be a continuous random variable and let:

$$dP = f(x)dx \qquad \ldots.(3.1)$$

be the probability that the variable has a value between x and $(x + dx)$; $f(x)$ is known as the probability distribution function of the variable x. The function $f(x)$ satisfies the relation:

$$\int f(x)dx = 1 \text{ (normalization)} \qquad \ldots.(3.1a)$$

where the integration is carried out over all allowed values of x.

Averages, denoted by a bar, may be calculated as soon as the distribution function is known; for example:

$$\overline{x^m} = \int x^m dP \; (m = 1, 2, \ldots) \qquad \ldots.(3.2)$$

where the integration extends over all allowed values of x. The most important averages are $\bar{x}$ and $\overline{x^2}$. If $f(x)$ is symmetrical in x, the averages of all odd powers of x are zero; if $\bar{x} \neq 0$ it is advisable to introduce $(x - \bar{x})$ as a new random variable.

If n is a discrete random variable which can only have integral positive values and if $P(n)$ is the probability that the value n occurs, then the integrations in equations (3.1a) and (3.2) have to be replaced by summations. Equation (3.1a) becomes:

$$\sum_n P(n) = 1 \qquad \ldots.(3.1b)$$

and equation (3.2) is replaced by:

$$\overline{n^m} = \sum_n n^m P(n) \quad (m = 1, 2, \ldots) \qquad \ldots.(3.2a)$$

For two continuous random variables x and y the probability that the one variable has a value between x and $(x + \mathrm{d}x)$ and the other variable a value between y and $(y + \mathrm{d}y)$ is:

$$\mathrm{d}P = f(x, y)\mathrm{d}x\,\mathrm{d}y \qquad \ldots.(3.3)$$

in analogy with equation (3.1). The function $f(x, y)$ is called the joint probability distribution function of the variables x and y. It satisfies the condition:

$$\iint f(x, y)\mathrm{d}x\,\mathrm{d}y = 1 \text{ (normalization)} \qquad \ldots.(3.3a)$$

where the integration extends over all allowed values of x and y.

Averages are defined in the same manner as for one variable; that is:

$$\overline{x^n y^m} = \iint x^n y^m f(x, y)\mathrm{d}x\,\mathrm{d}y \qquad \ldots.(3.4)$$

where the integration extends again over all allowed values of x and y. Usually $\bar{x} = \bar{y} = 0$, the most important averages are then $\overline{x^2}$, $\overline{y^2}$ and $\overline{xy}$. If $\overline{xy} = 0$, the random variables x and y are said to be uncorrelated; if $\overline{xy} \neq 0$ the random variables are said to be correlated and the quantity:

$$c = \frac{\overline{xy}}{\sqrt{\overline{x^2}.\overline{y^2}}} \qquad \ldots.(3.5)$$

known as the correlation coefficient, is introduced. It is easily shown that $-1 \leqq c \leqq 1$.

If x and y are discrete random variables instead of continuous ones, then the above integrations have to be replaced by summations. If two random variables x and y are partly correlated, then y can be split into a part ax, fully correlated with x, and a part z, uncorrelated with x, by writing:

$$y = ax + z \qquad \ldots.(3.6)$$

where $\bar{x} = \bar{y} = \bar{z} = 0$ and $\overline{xz} = 0$. If c is the correlation coefficient of the two quantities, then:

$$a = c\left(\frac{\overline{y^2}}{\overline{x^2}}\right)^{1/2}; \quad \overline{z^2} = \overline{y^2}\,(1 - c^2) \qquad \ldots.(3.6a)$$

A random process described by a time-dependent random variable $X(t)$ is 'stationary' if the distribution function $f(X)$ does not depend explicitly upon t. In such a case the averages are also independent of t, since:

$$\overline{g(X)} = \int_0^\infty g(X)f(X)\mathrm{d}X \qquad \ldots.(3.7)$$

does not explicitly contain t either.

The average thus taken is known as an 'ensemble average', that is, an average taken over a large number of identical systems subjected to independent fluctuations (ensemble). In stationary processes $\overline{g(X)}$ can also be obtained by taking the limit:

$$\overline{g(X)} = \lim_{T\to\infty} \frac{1}{T}\int_0^T g(X)\mathrm{d}t \qquad \ldots.(3.7a)$$

Both methods give identical results. In our calculations, an average always means an ensemble average; noise measurements, however, are made by taking time averages over a sufficiently long time, the averaging usually being done by the measuring instrument (the quadratic detector with time constant τ) itself.

In stationary random processes the quantity $\overline{X(t)X(t+s)}$, known as the 'autocorrelation function', is very important. It is, of course, independent of time and, in addition, has the properties:

(1) Except when $\overline{X(t)X(t+s)}$ is a delta function in s, the quantity is continuous, even if $X(t)$ is discontinuous.

(2) $\overline{X(t)X(t+s)} = \overline{X^2(t)}$ for $s = 0$.

(3) $\overline{X(t)X(t+s)}$ is symmetrical in s.

If $\overline{X(t)X(t+s)}$ is a delta function in s, the noise source is said to be a white source; one often tries to refer noise phenomena back to white source functions.

The autocorrelation coefficient $c(s)$ is defined as:

$$c(s) = \frac{\overline{X(t)X(t+s)}}{\overline{X^2(t)}} \qquad \ldots.(3.8)$$

and is also known as the normalized autocorrelation function [normalization means $c(s) = 1$ for $s = 0$].

The following discrete distribution functions of the discrete variable $n (n = 0, 1, 2, \ldots)$ are of interest in noise caused by carrier density fluctuations or by emission fluctuations.

(1) *The binominal distribution.* Let a certain event have the probability p of occurring in the form A and the probability $(1 - p)$ of occurring in the form B and let individual events be independent.

If the event occurs m times, then the probability that n of them occur in the form A is:

$$P_m(n) = \frac{m!}{n!\,(m-n)!}\,p^n(1-p)^{m-n} \qquad \ldots.(3.9)$$

and*

$$\bar{n} = mp; \quad \overline{(n-\bar{n})^2} = mp(1-p) \qquad \ldots.(3.9a)$$

(2) *The Poisson distribution.* Let individual events be independent and let them occur at random at the average rate $\bar{n}$. The probability that n events occur in a given unit time interval is then:

$$P(n) = \frac{(\bar{n})^n \exp(-\bar{n})}{n!} \qquad \ldots.(3.10)$$

and:

$$\overline{(n-\bar{n})^2} = \bar{n} \qquad \ldots.(3.10a)$$

(3) *The normal distribution.* Let events occur at an average (large) rate $\bar{n}$ and let $\overline{(n-\bar{n})^2} = \sigma^2$. Then the probability that n events occur as a given unit time interval is:

$$P(n) - \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(n-\bar{n})^2}{2\sigma^2}\right] \qquad \ldots.(3.11)$$

The binominal and the Poisson distributions reduce to the normal distribution for large values of n. The particular distribution in which $\sigma^2 = \bar{n}$ is known as the Gaussian distribution.

In the case of a continuous random variable having $\bar{x} = 0$ a distribution function of the form:

$$\mathrm{d}P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)\mathrm{d}x \qquad \ldots.(3.12)$$

is usually found where:

$$\sigma^2 = \overline{x^2} \qquad \ldots.(3.12a)$$

This is known as the normal law for a continuous random variable.

Fourier Analysis of Fluctuating Quantities

Let $X(t)$ be a random variable describing a stationary random

* According to equation (3.2a):

$$\bar{n} = \sum_{n=0}^{\infty} nP_m(n); \quad \overline{n^2} = \sum_{n=0}^{\infty} n^2P_m(n), \text{ etc.}$$

Generally:

$$\overline{(n-\bar{n})^2} = \overline{n^2} - \overline{2n\bar{n}} + (\bar{n})^2 = \overline{n^2} - (\bar{n})^2$$

process, then $\overline{X^2(t)}$ will be independent of time. This mean square value may be written as:

$$\overline{X^2(t)} = \int_0^\infty S(f)\mathrm{d}f \qquad \ldots.(3.13)$$

where $S(f)$ is the spectral intensity of the fluctuating quantity $X(t)$ It may be shown that:

$$S(f) = 4\int_0^\infty \overline{X(t)X(t+s)}\cos \omega s\mathrm{d}s \qquad \ldots.(3.14)$$

and, by inversion, that:

$$\overline{X(t)X(t+s)} = \int_0^\infty S(f)\cos \omega s\mathrm{d}f \qquad \ldots.(3.15)$$

Equation (3.14) is known as the Wiener–Khintchine theorem.

The normalized autocorrelation function, $c(s)$, may then be introduced as:

$$S(f) = 4\overline{X^2(t)}\int_0^\infty c(s)\cos \omega s\mathrm{d}s \qquad \ldots.(3.16)$$

so that the spectral intensity $S(f)$ is known if $\overline{X^2(t)}$ and $c(s)$ are known. Calculating $S(f)$ thus usually means determining $c(s)$ and $\overline{X^2(t)}$ and, having determined $S(f)$, it is possible to determine the autocorrelation function $\overline{X(t)X(t+s)}$ from equation (3.15).

The spectral intensity $S(f)$ has the following interesting property: suppose a random signal $X(t)$ is applied to the input of an arbitrary linear system, having a transfer function $g(f)$, and a signal $Y(t)$ taken from the system. If $S_x(f)$ and $S_y(f)$ are the corresponding spectral densities, then:*

$$S_y(f) = S_x(f)\,|\,g(f)\,|^2 \text{ and } \overline{Y^2(t)} = \int_0^\infty S_x(f)\,|\,g(f)\,|^2\mathrm{d}f \qquad \ldots.(3.17)$$

The quantity $S(f)$ is related to the Fourier amplitude of $X(t)$. Let $X(t)$ be developed into a Fourier series for the time interval $0 \leqq t \leqq T$ and let $x_n = c_n \cos(\omega_n t + \phi_n)$ be its Fourier coefficient

* This is the basis for the method of noise measurement (see Chapter 2, pp. 8–10). Let the noise amplifier be tuned at the frequency $f = f_0$ and let its bandwidth be so narrow that $S_x(f) \simeq S_x(f_0)$ for the pass band, then:

$$\overline{Y^2(t)} = S_x(f_0)\int_0^\infty g^2(f)\,\mathrm{d}f = S_x(f_0)g_0^2B_{\mathrm{eff}}$$

where $g_0 = g(f_0)$. The values of g_0 and the effective bandwidth B_{eff} can be determined with the help of a known signal.