NOTE ON FLUCTUATIONS AND THE SHOT EFFECT

By E. N. ROWLAND, B.A., Gonville and Caius College

[Communicated by Prof. R. H. Fowler]

[*Received* 3 May, *read* 16 May 1938]

1. The present author recently published* what he believed to be a proof and extension of Campbell's theorem on fluctuations. More recently a paper† has been published containing a criticism of this proof. It therefore seems desirable that some parts of the proof be amplified. Further explanation of the foundations of the theory seems required, in order that it shall not appear that the theory ignores the randomness of the events, and a more detailed account of the derivation of the mean value of functions of pairs of events (i.e. sums of the type $\Sigma_{r \neq s} h(t_r, t_s)$) is necessary.

Before discussing the postulates of the shot-effect theory it appears necessary to explain briefly the meanings of the required terms "probability, independence of events, and mean value" in the senses in which these terms are to be used in this particular discussion. We consider an arbitrarily large set of repetitions of an experiment and an associated observation. The concept of the infinite aggregate of possible observations is derived from the concept of this arbitrarily large set in the same way as the concept of the infinite aggregate of real numbers is derived from the concept of an arbitrarily large set of numbers. If the nature of the experiment is such that the proportion of a set of observations, in which a particular occurrence A is found, is almost certainly as near a number $P(A)$ as we choose, if the set is large enough, then the occurrence is said to have a probability $P(A)$ (a probability in the postulated aggregate of possible observations). Questions of probability now reduce to questions of proportion in a finite set. A probability must satisfy $0 \leqslant P(A) \leqslant 1$. If the occurrences A and B are that some single observable quantity X shall have a value in a set of numbers A or in a distinct set B, then $P(A+B) = P(A)+P(B)$. The aggregate of observations which contains an occurrence A is a subaggregate of the possible observations in the same sense as even integers are a special subaggregate of real numbers. We can deduce that the probability $P(A \text{ and } B)$ of finding an occurrence A and an occurrence B together is the probability $P(A)$ of A multiplied by the probability $P_A(B)$ of finding B in the subaggregate of observations which contain an occurrence A, provided that both $P(A)$ and $P_A(B)$ exist. This last probability may or

* *Proc. Cambridge Phil. Soc.* 32 (1936), 580–97.

† "The shot effect for showers", by J. M. Whittaker, *Proc. Cambridge Phil. Soc.* 33 (1937), 451–8 (456).

https://doi.org/10.1017/S0305004100020259 Published online by Cambridge University Press

may not be the probability $P(B)$ of B in the whole aggregate. If it is, then the occurrences are said to be independent and $P(A \text{ and } B) = P(A)\,P(B)$. Lastly, to establish a mean value of some observable quantity X we consider the mean value $\mathscr{M}_k X$ of X in a set of k observations. If it can be shown that, when k is sufficiently large, $\mathscr{M}_k X$ will almost certainly be arbitrarily near some quantity, then the mean value $\mathscr{M}X$ of X is defined to be this last quantity. It follows that $\mathscr{M}(X+Y) = \mathscr{M}X + \mathscr{M}Y$ and that if X has probabilities $P(x_1)$, $P(x_2)$, ..., of being x_1, x_2, ..., and if these numbers are bounded and form a complete set of possible values of X, then

$$\mathscr{M}X = x_1 P(x_1) + x_2 P(x_2) + \ldots. \tag{1}$$

If these numbers are not bounded this last expression need not converge and the basic assumptions have to be extended if the theorem (1) is to be proved.

2. The postulates of our discussion of phenomena like the shot effect without secondary emission or space-charge effects are that the repetitions of an experiment and the associated aggregate of possible observations yield a probability $N\,dt + O(dt)^2$ that a small interval $(t, t+dt)$ contains a single event and a probability $O(N\,dt)^m$ that it contains m events, that events in separate intervals are independent, and that there is some finite limit to the number of events an interval may contain. The last condition allows us to apply the above theorem (1) and we deduce that if X is the number of events in $(t, t+dt)$, then

$$\mathscr{M}X = N\,dt + O(dt)^2.$$

Since mean values are additive, it follows that $\mathscr{M}X$ must be a linear function of dt, i.e. $\mathscr{M}X = N\,dt$. From the property of the independence of events in different intervals (i.e. of the randomness of the events) we deduce that the probability of a pair of events, one in $(t, t+dt)$ and the other in the separate interval (t_1, t_1+dt_1), is $N^2 dt\,dt_1 + O(dt+dt_1)^2$. If X is the number of pairs with one event in each of the intervals we deduce that $\mathscr{M}X = N^2 dt\,dt_1 + O(dt+dt_1)^2$. Since $\mathscr{M}X$ must be bilinear in dt and dt_1 this reduces to $\mathscr{M}X = N^2 dt\,dt_1$. It will be observed that the randomness of the events in a single possible observation is reduced to the evenness of the mean distribution. If, for example, the events were not random but nearly evenly spaced, then the mean distribution of pairs of events would have a peak for pairs at the distance of the spacing and negligible density for pairs closer together. Any deviation of the mean densities of single events and of pairs from the values N and N^2 corresponds to some ordering of the events. The only other assumption we require is that the functions dealt with are bounded and have proper Riemann integrals.

The process which shows that

$$\mathscr{M}\Sigma_r h(t-t_r) = \int_{-T_0}^{T_1} h(\tau)\,N\,d\tau \tag{2}$$

has been given in full in the previous paper and it was stated that for a function

https://doi.org/10.1017/S0305004100020259 Published online by Cambridge University Press

such as $\Sigma_{r\neq s} h(t_r, t_s)$ of pairs of events a practically identical argument would show that

$$\mathscr{M}\, \Sigma_{r\neq s} h(t_r, t_s) = \int_{-T_0, -T_0}^{T_1, T_1} h(\tau, \tau')\, d\tau d\tau'. \tag{3}$$

J. M. Whittaker appears to maintain that the argument would show that $\mathscr{M}\, \Sigma_{r,s} h(t_r, t_s)$ is equal to this integral, and consequently it is desirable to discuss the proof in detail. Now the quantity $\Sigma_{r,s} h(t_r, t_s)$ consists of a contribution $\Sigma_{r\neq s} h(t_r, t_s)$ from pairs of events and of a contribution $\Sigma_r h(t_r, t_r)$ from separate events. We divide the interval $(-T_0, T_1)$ into subintervals $(\tau_\mu, \tau_{\mu+1})$ of length $d\tau_\mu$. Almost all the pairs of events are situated with members in separate subintervals, though a small proportion are situated with both members in the same subinterval. Then

$$\Sigma_{r\neq s} h(t_r, t_s) = \Sigma_{\mu,\nu} \Sigma_{t_r \text{ in } (\tau_\mu, \tau_{\mu+1})}^{t_s \text{ in } (\tau_\nu, \tau_{\nu+1})} h(t_r, t_s), \tag{4}$$

the summation extending over all pairs of intervals and including the terms with $\mu = \nu$ but not $r = s$, for any interval which contains more than one event. The mean of this quantity in a set of k repetitions is $\Sigma_{\mu,\nu} \mathscr{M}_k \Sigma_{t_r \text{ in } (\tau_\mu, \tau_{\mu+1})}^{t_s \text{ in } (\tau_\nu, \tau_{\nu+1})} h(t_r, t_s)$. Now consider the contribution to this from a single pair of subintervals $(\tau_\mu, \tau_{\mu+1})$ and $(\tau_\nu, \tau_{\nu+1})$. A pair of events falling in this pair of intervals yields nearly* $h(\tau_\mu, \tau_\nu)$. As has been stated above, the mean number of pairs with a member in each of these intervals is $N^2 d\tau_\mu d\tau_\nu$ and so the mean number for the set is almost certainly nearly $N^2 d\tau_\mu d\tau_\nu$. Hence the contribution from this pair of intervals is almost certainly arbitrarily near $h(\tau_\mu, \tau_\nu) N^2 d\tau_\mu d\tau_\nu$. Suppose also that the mean number of pairs of events in a single interval of length dt is $N^2 dt^2$, so that terms in the summation with $\mu = \nu$ have the same form as the rest: it is shown below that our previous hypothesis that the mean number is merely $O(dt)^2$ is sufficient. Then, for sufficiently large k, $\mathscr{M}_k \Sigma_{r\neq s} h(t_r, t_s)$ is almost certainly arbitrarily near $\Sigma_{\mu,\nu} h(\tau_\mu, \tau_\nu) N^2 d\tau_\mu d\tau_\nu$. For, if it differs from this by an appreciable amount, then at least one part must differ by at least this amount divided by the number of pairs of subintervals and, for sufficiently large k, this is almost certain not to happen. Hence, if the division τ_μ was originally chosen fine enough, the last expression is arbitrarily near the above integral and we conclude that the mean value of $\Sigma_{r\neq s} h(t_r, t_s)$ for the set is almost certainly arbitrarily near the above integral. This proves the proposition (3), except that we have still to examine the contribution of pairs of events occurring in the same subinterval. With the hypothesis that the mean number of pairs in a subinterval is $N^2 dt^2$ this contribution is almost certainly nearly $\Sigma_\mu h(\tau_\mu, \tau_\mu) N^2 d\tau_\mu^2$ and, with a fine enough division,

* The following argument can readily be put into formal mathematical language as was shown in the previous paper with $\mathscr{M}\, \Sigma_r h(t_r)$. It is regretted that a term $\frac{1}{2}\epsilon(\min d\tau_\nu/(T_0 + T_1))$ was there inverted. The way to avoid the difficulty that the above statement about the yield of a pair of events is only an approximation and does not hold at the discontinuities is there shown by the use of $\underline{B}$ and $\overline{B}$, here omitted for simplicity.

https://doi.org/10.1017/S0305004100020259 Published online by Cambridge University Press

this is negligible. It would not matter if such pairs were neglected altogether or if the mean number of pairs in a subinterval were assumed to be merely $O(d\tau_\mu)^2$ instead of $N^2 d\tau_\mu^2$.

The above argument shows that the mean value, for the set, of $\Sigma_{r\neq s} h(t_r, t_s)$ is almost certainly nearly the above integral. To find what $\Sigma_{r,s} h(t_r, t_s)$ will probably be we have to add the probable value of $\Sigma_r h(t_r, t_s)$. A single subinterval $(\tau_\mu, \tau_{\mu+1})$ will probably contain on the average $N d\tau_\mu$ events and will probably contribute $h(\tau_\mu, \tau_\mu) N d\tau_\mu$ to $\mathscr{M}_k \Sigma_r h(t_r, t_r)$. This, unlike the $h(\tau_\mu, \tau_\mu) N^2 d\tau_\mu^2$ above, cannot be neglected, for its sum does not vanish as the division becomes finer. The sum $\Sigma_r h(t_r, t_r)$ is of the same form as the sum $\Sigma_r h(t_r)$ previously discussed, and so it is not necessary to repeat the rest of the argument which shows that its mean value for the set is almost certainly nearly $\int_{-T_0}^{T_1} h(\tau, \tau) N d\tau$. Addition of the two parts then shows that the mean value, for the set, of $\Sigma_{r,s} h(t_r, t_s)$ is almost certainly nearly the sum of the two integrals.

3. *Comparison with Whittaker's proof.* A further justification of the validity of the above proof can be obtained by showing that its basic algebraic operations are not very different from those of Whittaker's proof. A slight modification, which Whittaker accepts, puts his proof into the form

$$\theta^2 = \Sigma_{r,s} y_r y_s f(rL) f(sL) = \Sigma_r y_r^2 f(rL) + \Sigma_{r\neq s} y_r y_s f(rL) f(sL),$$

$$\overline{y_r^2} = \nu L + \mu^2 L^2, \quad \text{from the line above (20)},$$

$$\overline{y_r y_s} = \mu^2 L^2, \quad (r \neq s), \quad \text{from (31)},$$

$$\overline{\theta^2} = \Sigma_r (\nu L + \mu^2 L^2) f^2(rL) + \Sigma_{r\neq s} \mu^2 L^2 f(rL) f(sL)$$
$$= \Sigma_r \nu L f^2(rL) + \Sigma_{r,s} \mu^2 L^2 f(rL) f(sL).$$

When this is translated into our notation and reduced to the case of one type of event it becomes, if we use y_μ for the number of events in $(\tau_\mu, \tau_{\mu+1})$,

$$\vartheta^2 \simeq \Sigma_\mu y_\mu^2 h(\tau_\mu, \tau_\mu) + \Sigma_{\mu\neq\nu} y_\mu y_\nu h(\tau_\mu, \tau_\nu),$$

$$\overline{y_\mu^2} = N d\tau_\mu + O(d\tau_\mu)^2, \quad \overline{y_\mu y_\nu} = N^2 d\tau_\mu d\tau_\nu, \quad (\mu \neq \nu),$$

$$\overline{\vartheta^2} \simeq \Sigma_\mu (N d\tau_\mu + O(d\tau_\mu)^2) h(\tau_\mu \tau_\mu) + \Sigma_{\mu\neq\nu} N^2 d\tau_\mu d\tau_\nu h(\tau_\mu \tau_\nu)$$
$$\simeq \int_{-T_0}^{T_1} h(\tau, \tau) N d\tau + \int_{-T_0, -T_0}^{T_1, T_1} h(\tau, \tau') N^2 d\tau d\tau'.$$

The establishing of the values of $\overline{y_\mu^2}$ and $\overline{y_\mu y_\nu}$ for this simplified case does not require Whittaker's elaborate analysis. The equation for $\overline{y_\mu^2}$ is valid since y_μ has a probability $N d\tau_\mu + O(d\tau_\mu)^2$ of being 1 and $O(N d\tau_\mu)^m$ of being m. The equation for $\overline{y_\mu y_\nu}$ is valid since y_μ and y_ν are independent and entirely uncorrelated quantities. In the equations for $\overline{\vartheta^2}$ the terms of $\Sigma_\mu h(\tau_\mu, \tau_\mu) O(d\tau_\mu)^2$ are collected with

https://doi.org/10.1017/S0305004100020259 Published online by Cambridge University Press

those of $\Sigma_\mu h(\tau_\mu \tau_\mu) N d\tau_\mu$ because they come from the contribution of single intervals, while the remaining terms come from pairs of intervals. We collect the terms of $\Sigma_\mu h(\tau_\mu, \tau_\mu) O(d\tau_\mu)^2$ with the terms of the $\Sigma_{\mu \neq \nu}$ because they arise from pairs of events even though the pairs are in single intervals. Otherwise there is no difference between this form of Whittaker's proof and the algebra of ours.

The main difference between the author's proof of the theorem and Whittaker's is that the author does attempt to prove the existence of the mean value, or at least to satisfy our definition of a mean value. This is the reason for the extra elaboration of the above proof. Whittaker appears to assume the existence. He gives his results in the form

$$\bar{\theta} = \mu \sum_{r=1}^{k} Lf(rL)$$

which, in our notation, with N_p for the probable density of arrival of groups of p secondary electrons, is

$$\bar{\vartheta} = (\Sigma_p p N_p) \Sigma_\nu f(\tau_\nu) d\tau_\nu.$$

The author believes that this formula must be regarded as an approximation for an intuitive mean value. Whether the approximation converges or not as the division becomes finer appears to require a discussion with an upper and a lower limit fixed for $\bar{\theta}$ as in our discussion with $\bar{B}$ and $\underline{B}$ in the previous paper. It appeared in that discussion, that if $h(t)$ had not a Riemann integral the process of taking a finer division of the interval would not yield convergence of the approximations*. If it can be shown that, with suitable assumptions, the approximation does converge, then presumably the assumption of the existence of a mean value or the definition of the mean value as this limit would be justified. This method, however, does not bring to light the valuable fact that the mean for a sufficiently large number of repetitions of an observation is almost certainly nearly the mean, a fact which leads immediately to the conclusion that $T^{-1}\int_{-T_0}^{T_0} \vartheta(t)\, dt$, etc. may be taken as physically good estimates of $\bar{\vartheta}$, etc. when T is large.

4. *Corrections.* The author regrets that in both the paper cited and in a continuation† the modulus signs have been dropped from $|G^2|$ in various integrals. The amplification function G must in general be complex in order to allow for change of phase on amplification. Further, $G(-x)$ must be the complex conjugate of $G(x)$ because waves containing the time factor $e^{-i\nu t}$ are retarded as

* The process can be modified by introducing a division of the range of h as well as of the interval $(-T_0, T_1)$ and it can be shown that existence of the Lebesgue integral ensures convergence of the corresponding approximations. Campbell's theorem is therefore true with Lebesgue integrals. The existence of some integral is, however, a necessary postulate for the convergence.

† "The shot effect, II", *Proc. Cambridge Phil. Soc.* 33 (1937), 344–58.

https://doi.org/10.1017/S0305004100020259 Published online by Cambridge University Press

much as waves containing e^{ivt} are advanced in phase. This is also the necessary and sufficient condition that $f(t)$ shall be real when $V(t)$ is real. The results should therefore read

$$\overline{\vartheta^2} \simeq \frac{1}{\pi} jeR^2 \int_0^\infty |\, G^2(x)\,|\, dx, \text{ etc.} \qquad \text{(T44)}$$

Further, it appears that the calculation of $P(t-t_1)$ for (17) in the second paper is not correct. It should give

$$P(t-t_1) = \bar{N}^2 - \frac{\bar{N}}{C\rho} \frac{2\rho+R}{2(\rho+R)} \exp\left[-(t-t_1)\left(\frac{1}{CR}+\frac{1}{C\rho}\right)\right],$$

and the process of calculating the shot effect then gives Moullin's formula

$$\overline{\vartheta^2} - \bar{\vartheta}^2 \simeq \bar{N}\left(1+\frac{R}{\rho}\right)^{-2} \frac{1}{\pi} e^2R^2 \int_0^\infty |\, G^2\,|\, dx$$

for amplification of frequencies less than $1/CR$, and

$$\overline{\vartheta^2} - \bar{\vartheta}^2 = \left(1+\frac{R}{\rho}\right)^{-1} \tfrac{1}{2}\bar{N}e^2\frac{R}{C}$$

for the total fluctuation at the anode. It is hoped to publish a proof of this result in due course. These formulae are calculated on the assumption that the arrival of electrons at the anode is random but that the probable densities vary with the anode voltage in the way required to give the correct mean currents with various voltages. Recent experimental work by F. C. Williams* has shown that a further factor dependent on the anode current has to be introduced to describe the observed shot noise. It follows that the space charge itself introduces some correlation between neighbouring arrivals of electrons and that $P(t-t_1)$ requires some further modification which is difficult to estimate from the present concepts of the space charge.

* "Fluctuation voltage in diodes and in multielectrode valves", *J. Inst. Elec. Eng.* 79 (1936), No. 477, pp. 349–60.

https://doi.org/10.1017/S0305004100020259 Published online by Cambridge University Press