A mathematical comparison of simple models of Johnson noise and shot noise

2000 J. Phys.: Condens. Matter 12 4195

(http://iopscience.iop.org/0953-8984/12/18/305)

J. Phys.: Condens. Matter **12** (2000) 4195–4205. Printed in the UK PII: S0953-8984(00)10562-4

A mathematical comparison of simple models of Johnson noise and shot noise

Daniel T Gillespie
Research Department (Code 4T4100D), Naval Air Warfare Center, China Lake, CA 93555, USA
E-mail: gillespiedt@navair.navy.mil

Received 4 January 2000

Abstract. Simple physical models of Johnson noise and shot noise are compared within the mathematical framework of continuous Markov process theory. The comparison may help illuminate differences or misconceptions that might be fuelling the current debate over whether Johnson noise and shot noise can be subsumed under a single, unified theory.

1. Introduction

Johnson noise and shot noise are well documented manifestations of stochastically moving electrons. Although these two kinds of electrical noise were discovered at different times [1, 2] and initially explained in different ways [3, 4], effort has recently been devoted to establishing a unified physical theory for them [5, 6]. But this attempt to unify Johnson noise and shot noise has generated controversy [7, 8], and experts are currently divided over whether such a unification is possible.

Scientific controversy is often rooted either in hidden differences in fundamental assumptions and definitions, or else in a lack of awareness of all the logical implications of those assumptions and definitions. With that in mind, the aim of this paper will not be to finally resolve the present controversy, but merely to illuminate the arena of its debate. More specifically, the aim here will be to derive and compare the main predictions of the *simplest theoretical models* of Johnson noise and shot noise *within a common mathematical framework*. That framework, continuous Markov process theory, is briefly reviewed in section 2. The fact that Johnson noise can be thus analysed is well known, and its analysis is summarized in section 3. The fact that shot noise too can be succinctly analysed using continuous Markov process theory is apparently less well known, so its analysis in section 4 is drawn more carefully. The concluding section 5 compares and contrasts the findings for the two models.

To broadly anticipate those findings, there seem to be substantial mathematical differences between Johnson noise and shot noise. Proponents of a unified theory of electrical noise might choose to argue that these differences can be bridged through an amended analysis, or they might see fit to argue that the differences are irrelevant because at least one of the specific physical models assumed here for the noises is deficient. But either response could help move the debate closer to a resolution.

2. A short review of continuous Markov process theory

Before we try to analyse Johnson noise and shot noise in terms of continuous Markov process theory, it is appropriate to review the relevant features of that theory. Details and elaborations on the following review may be found in [9].

0953-8984/00/184195+11$30.00

For our purposes here, a *continuous Markov process* can be defined as a function X of time t that evolves according to the following rule: given the value $X(t)$ of X at time t, the value of X at any infinitesimally later time $t + \mathrm{d}t$ can be computed from the formula

$$X(t + \mathrm{d}t) = X(t) + A(X(t), t)\,\mathrm{d}t + D^{1/2}(X(t), t)N(t)(\mathrm{d}t)^{1/2}. \tag{1}$$

This formula is called the *standard-form Langevin equation*, and as we shall see shortly it is not as whimsically arbitrary as it might at first appear. In this equation, A and D are given smooth functions of their two arguments, with D non-negative; $\mathrm{d}t$ is a *real variable* confined to the interval $[0, \varepsilon]$, where $\varepsilon > 0$ is arbitrarily small; and $N(t)$ is a 'temporally uncorrelated unit normal random variable'. Expanding on the last, if $\mathcal{N}(m, \sigma^2)$ denotes the normal (Gaussian) random variable with mean m and variance σ^2, then $N(t) = \mathcal{N}(0, 1)$, with $N(t)$ and $N(t')$ statistically independent whenever $t \neq t'$.

The function A is called the *drift function*; it evidently enters equation (1) in such a way as to produce a deterministic drift of the process in time. The function D is called the *diffusion function*; since its square root is multiplied in equation (1) by the zero-mean, unit-variance normal random variable $N(t)$, it gives rise to random fluctuations in the process about its deterministic drift. The fact that $(\mathrm{d}t)^{1/2} \gg \mathrm{d}t$ means that the diffusion term in equation (1) will usually be much larger than the drift term; however, over a *succession* of many $\mathrm{d}t$ increments, the random sign changes in the values of $N(t)$ render the *cumulative effect* of the diffusion term comparable to that of the drift term, provided the magnitudes of A and $D^{1/2}$ are comparable.

Two important properties of the process X defined by the infinitesimal updating formula (1) are immediately obvious from the form of that formula: first, X is *continuous*, because equation (1) implies that $X(t + \mathrm{d}t) \to X(t)$ as $\mathrm{d}t \to 0$, and second, X is *memoryless* or *Markovian*, because the recipe prescribed by equation (1) for computing $X(t + \mathrm{d}t)$ from $X(t)$ requires no knowledge of any values of X *before* time t.

An important property of equation (1) that is *not* immediately obvious is its *unique self-consistency*: if we advance from t to $t + \mathrm{d}t$ in *two* steps, say first from t to $t + \frac{1}{2}\mathrm{d}t$ and then from $t + \frac{1}{2}\mathrm{d}t$ to $t + \mathrm{d}t$, the corresponding *two* applications of equation (1) will give the *same final result* (statistically and to lowest order in $\mathrm{d}t$) as a *single* application of equation (1) in its t to $t + \mathrm{d}t$ form. Furthermore, any alterations to the *form* of equation (1)—such as taking $N(t)$ to be non-normal or changing the exponents of $\mathrm{d}t$ in either term on the right-hand side—will *spoil* this self-consistency. So, far from being whimsically arbitrary, the infinitesimal updating formula (1) is in fact the *only self-consistent* formula possible *if* the process X so defined is to be both *continuous* and *memoryless*.

A result that should be mentioned for the sake of completeness (we will not need it in the following) is that the probability density function of $X(t)$, given that $X(t_0) = x_0$ for $t_0 < t$, satisfies the partial differential equation

$$\frac{\partial P(x, t|x_0, t_0)}{\partial t} = -\frac{\partial}{\partial x}[A(x, t)P(x, t|x_0, t_0)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[D(x, t)P(x, t|x_0, t_0)]. \tag{2}$$

This is called the (*forward*) *Fokker–Planck equation*, and its mathematical equivalence to the Langevin equation (1) is one of the key theorems of continuous Markov process theory. Equations (1) and (2) make it clear that a continuous Markov process $X(t)$ is completely specified by the forms of its A and D functions.

If we write the standard-form Langevin equation (1) in the form

$$\frac{X(t + \mathrm{d}t) - X(t)}{\mathrm{d}t} = A(X(t), t) + D^{1/2}(X(t), t)N(t)(\mathrm{d}t)^{-1/2} \qquad (\mathrm{d}t \neq 0) \tag{3}$$

we can see that, in the special case $D \equiv 0$, we can take the limit $\mathrm{d}t \to 0$ and obtain

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = A(X(t), t) \qquad (D \equiv 0). \tag{4}$$

This ordinary differential equation defines a *deterministic* process, and it shows that 'noiseless' continuous Markov processes are just the familiar stuff of ordinary calculus: if $D \equiv 0$, then $X(t)$ is not only *continuous* but also *differentiable*. But if $D > 0$, the last term in equation (3) is evidently ill behaved in the limit $\mathrm{d}t \to 0$; hence, *genuinely stochastic* continuous Markov processes are *not* differentiable.

In spite of the last mentioned fact, it has become customary to 'pretend' that $\mathrm{d}X(t)/\mathrm{d}t$ exists even in the case $D > 0$. This is accomplished through the following mathematical ruse: since the normal random variable has the property that

$$a + b\mathcal{N}(m, \sigma^2) = \mathcal{N}(a + bm, b^2\sigma^2) \tag{5}$$

then

$$N(t)(\mathrm{d}t)^{-1/2} = (\mathrm{d}t)^{-1/2}\mathcal{N}(0, 1) = \mathcal{N}(0, 1/\mathrm{d}t).$$

So, if we formally *define*

$$\Gamma(t) \equiv \lim_{\mathrm{d}t \to 0} \mathcal{N}(0, 1/\mathrm{d}t) \tag{6}$$

then the $\mathrm{d}t \to 0$ limit of equation (3) yields

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = A(X(t), t) + D^{1/2}(X(t), t)\Gamma(t). \tag{7}$$

Equation (7) is called the *white-noise form Langevin equation*, and it may be regarded as a *formal equivalent* of the standard-form Langevin equation (1). The quantity $\Gamma(t)$ defined in equation (6) is called *temporally uncorrelated Gaussian white noise*; it satisfies the two averaging relations

$$\langle \Gamma(t) \rangle = 0 \tag{8a}$$

$$\langle \Gamma(t)\Gamma(t') \rangle = \delta(t - t') \tag{8b}$$

where δ is the Dirac delta function. Equation (8*a*) holds because the normal random variable on the right side of equation (6) has zero mean. The $t \neq t'$ case of equation (8*b*) holds because $\Gamma(t)$ is temporally uncorrelated, which follows of course from the fact that $N(t)$ in equation (1) is temporally uncorrelated. And finally, the $t = t'$ case of equation (8*b*), which asserts that $\mathrm{var}\{\Gamma(t)\} = \delta(0)$, follows formally from the implication of definition (6) that $\mathrm{var}\{\Gamma(t)\} = 1/(\mathrm{d}t \to 0)$ since $\delta(0)\,\mathrm{d}t = 1$ in the limit $\mathrm{d}t \to 0$.

Although the *derivative* of a continuous Markov process $X(t)$ is of questionable legitimacy, its *integral Y(t)* can be solidly defined by $\mathrm{d}Y(t)/\mathrm{d}t = X(t)$, or equivalently

$$Y(t + \mathrm{d}t) = Y(t) + X(t)\,\mathrm{d}t. \tag{9}$$

But since this defining equation does not have the canonical form (1)—because its right-hand side involves a process *other* than Y—then the integral of a continuous Markov process is *not itself* a Markov process. (However it turns out that $X(t)$ and $Y(t)$ *together* constitute a *bivariate* Markov process.)

In the discussion of Johnson noise and shot noise in the following, we shall encounter two different kinds of continuous Markov process. One of these is the *Wiener* process; it is defined by the drift and diffusion functions

$$A(x, t) = a \text{ and } D(x, t) = c \tag{10a}$$

where a is any real constant and c is any *positive* real constant. Using either equation (1) or equation (2), it can be shown that for the initial condition $X(t_0) = x_0$, the Wiener process $X(t)$ for any $t > t_0$ is the normal random variable

$$X(t) = \mathcal{N}(x_0 + a(t - t_0), c(t - t_0)) \qquad \text{(Wiener process).} \tag{10b}$$

The other type of continuous Markov process that will be of interest to us here is the *Ornstein–Uhlenbeck* process. It is defined by the drift and diffusion functions

$$A(x,t) = -\frac{x}{\tau} \text{ and } D(x,t) = c \tag{11a}$$

where τ and c are any *positive* real constants. Using either equation (1) or equation (2), it can be shown that for the initial condition $X(t_0) = x_0$, the Ornstein–Uhlenbeck process $X(t)$ for any $t > t_0$ is the normal random variable

$$X(t) = \mathcal{N}\left(x_0 \mathrm{e}^{-(t-t_0)/\tau}, \frac{c\tau}{2}(1 - \mathrm{e}^{-2(t-t_0)/\tau})\right) \qquad \text{(Ornstein–Uhlenbeck process).} \tag{11b}$$

We recall that any stochastic process $Z(t)$ (Markovian or not) that is *stationary*, in the sense that its first two moments are both constant in time, will have an auto-covariance function

$$\langle Z(t)Z(t+t')\rangle \equiv C_Z(t') \qquad (t' \geqslant 0) \tag{12}$$

that is independent of t. Moreover, by the Wiener–Khintchine theorem, the positive-frequency Fourier transform of this auto-covariance function, namely

$$S_Z(\nu) \equiv 4\int_0^\infty C_Z(t')\cos(2\pi\nu t')\,\mathrm{d}t' \qquad (\nu \geqslant 0) \tag{13}$$

will be such that $S_Z(\nu)\,\mathrm{d}\nu$ measures the portion of the (constant) second moment $\langle Z^2(t)\rangle$ associated with fluctuations of $Z(t)$ in the frequency range $[\nu, \nu + \mathrm{d}\nu]$. $S_Z(\nu)$ is called the *spectral density function* of $Z(t)$, although it really describes the frequency spectrum of $\langle Z^2(t)\rangle$ rather than $Z(t)$.

The stochastic process $a\Gamma(t)$ evidently qualifies as a stationary process (its mean is zero and its second moment is infinite), and using equations (8*b*), (12) and (13) it is easy to show that its spectral density function is

$$S_{a\Gamma}(\nu) = 2a^2 \qquad (\nu \geqslant 0). \tag{14}$$

The designation *white* noise for $\Gamma(t)$ derives of course from the fact that $S_\Gamma(\nu)$ is independent of ν. The fact that the integral of equation (14) over all frequencies is infinite is in keeping with the infinite second moment of $\Gamma(t)$.

The Wiener process of equations (10) is evidently *not* a stationary process, so it does not have a spectral density function (at least in the strict sense defined above). The Ornstein–Uhlenbeck process of equations (11) is *asymptotically* stationary, in the sense that its $t_0 \to -\infty$ limit is stationary:

$$X^*(t) \equiv \lim_{t_0 \to -\infty} X(t) = \mathcal{N}(0, c\tau/2) \qquad \text{(asymptotic O–U process).} \tag{15}$$

It can be shown through a moderately lengthy analysis that the spectral density function of the asymptotic Ornstein–Uhlenbeck process is

$$S_{X^*}(\nu) = \frac{2c\tau^2}{1 + (2\pi\tau\nu)^2} \qquad (\nu \geqslant 0). \tag{16}$$

It is easy to verify that the integral of equation (16) over all $\nu \geqslant 0$ gives the expected result $c\tau/2 = \langle X^{*2}(t)\rangle$.

3. Johnson noise

Johnson noise manifests itself in the temporally fluctuating electrical current $I_J(t)$ in a passive wire loop of resistance R and self-inductance L at absolute temperature T. As is shown in [9], a concise quantitative theory of Johnson noise can be built on the *assumption* that the many

complex interactions between the conducting electrons and the thermally vibrating atomic lattice of the wire effectively give rise to a *thermal emf* in the loop that is mathematically expressible as a sum of *two terms*: a *dissipative emf*, with the ohmic form $-RI_J(t)$; and a randomly fluctuating *Johnson emf* $V_J(t)$, which has zero mean and is independent of $I_J(t')$ for all $t' \leqslant t$. Integrating the electric potential once around the loop therefore gives the circuit equation

$$-L\frac{\mathrm{d}I_J(t)}{\mathrm{d}t} + [-RI_J(t) + V_J(t)] = 0. \tag{17a}$$

Rewriting this equation as

$$\frac{\mathrm{d}I_J(t)}{\mathrm{d}t} = -\frac{I_J(t)}{L/R} + \frac{1}{L}V_J(t) \tag{17b}$$

and then comparing with equations (7) and (11a), we see that it would define $I_J(t)$ as an Ornstein–Uhlenbeck process with

$$\tau = L/R \tag{18}$$

if if were so that

$$V_J(t) = Lc^{1/2}\Gamma(t) \tag{19}$$

where $\Gamma(t)$ is temporally uncorrelated Gaussian white noise and c is any positive constant. Our earlier assumption that $V_J(t)$ randomly fluctuates with zero mean and is independent of $I_J(t')$ for $t' \leqslant t$ is clearly *compatible* with equation (19). But more than that, our implicit assumption that equation (17) *self-consistently* governs the time evolution of $I_J(t)$ essentially *demands* that $I_J(t)$ be a *continuous Markov* process, and the Ornstein–Uhlenbeck process is the *only* such process whose form is consistent with our initial assumptions regarding the thermal emf. Equation (19) is therefore valid, but it ought not be counted as an additional assumption.

It remains only to determine the constant c, and this can be done as follows. Let us denote the equilibrium or steady-state current in the wire loop by $I_J^*(t)$:

$$I_J^*(t) \equiv \lim_{t_0 \to -\infty} I_J(t). \tag{20}$$

The assumption that the loop be in 'thermal equilibrium at absolute temperature T' means that the average equilibrium energy of the current in the loop must obey the equipartition theorem of thermodynamics; thus, $\langle \frac{1}{2}LI_J^{*2}(t)\rangle = \frac{1}{2}kT$, or

$$\langle I_J^{*2}(t)\rangle = kT/L \tag{21}$$

where k is Boltzmann's constant. But since $I_J(t)$ is an Ornstein–Uhlenbeck process, then we also have the asymptotic Ornstein–Uhlenbeck result (15):

$$\langle I_J^{*2}(t)\rangle = c\tau/2. \tag{22}$$

Combining the last two equations, and invoking the formula (18) for τ, we conclude that

$$c = \frac{2kTR}{L^2}. \tag{23}$$

The Johnson noise current $I_J(t)$ is thus completely characterized: it is an Ornstein–Uhlenbeck process whose defining constants τ and c are given by equations (18) and (23). Inserting those expressions for τ and c into the generic Ornstein–Uhlenbeck formulae (11b) and (16), we find that for the initial condition $I_J(t_0) = I_0$ the process $I_J(t)$ for any $t > t_0$ is the normal random variable

$$I_J(t) = \mathcal{N}\left(I_0 \exp\left(-\frac{(t-t_0)}{L/R}\right), \frac{kT}{L}\left(1 - \exp\left(-\frac{2(t-t_0)}{L/R}\right)\right)\right) \tag{24}$$

and also that the spectral density function of the equilibrium Johnson noise current is

$$S_{I_J^*}(\nu) = \frac{4kT}{R}\left(\frac{1}{1+(2\pi L\nu/R)^2}\right). \tag{25}$$

Another important result emerges upon substituting equation (23) into equation (19):

$$V_J(t) = (2kTR)^{1/2}\Gamma(t). \tag{26}$$

This is essentially the *fluctuation–dissipation theorem*†. It asserts that the assumed two components of the thermal emf, namely the *fluctuating* Johnson emf $V_J(t)$ and the *dissipative* ohmic emf $-RI_J(t)$, are *not independent* of each other, but are monotonically related through the loop's resistance R. At any fixed temperature, increasing (or decreasing) either one of those two component emfs will necessarily entail increasing (or decreasing) the other. There can be no fluctuation without dissipation, and conversely. The physical reason for this is that, even though we can *mathematically* decompose the thermal emf into two terms, the microscale processes that are ultimately responsible for the thermal emf cannot neatly be separated into one kind of process that produces *only* random fluctuations and another kind that produces *only* dissipation. We should also note that, by combining equation (26) with the result (14), we can immediately conclude that the spectral density function of the Johnson emf $V_J(t)$ is the constant $4kTR$; this result is often referred to as *Nyquist's theorem*.

Equations (24), (25) and (26) encapsulate the key properties of Johnson noise that follow, within the framework of continuous Markov process theory, from the simple physical assumptions set forth in the first paragraph of this section. Those assumptions—the existence of a 'thermal emf' that is mathematically expressible as the sum of a dissipative ohmic term $-RI_J(t)$ and a zero-mean randomly fluctuating term $V_J(t)$—are rather 'macroscopic' in character. One might wonder is a different picture of Johnson noise would emerge if we adopted a more 'microscopic' view [10], focusing for example on the actual collisions of the conduction electrons with the thermally vibrating atomic lattice of the wire. While not claiming to answer this question definitively, the present writer will offer his *opinion* that any such microscopic modelling of Johnson noise will lead to results that merely elaborate, but in no way contradict, the results of our present analysis.

This opinion is based on results in earlier investigations of the closely related phenomenon of *Brownian motion*. One way of analysing the velocity $U_B(t)$ of a Brownian particle is to proceed from the 'macroscopic' assumption that the total force exerted on the particle by its surrounding fluid can be written as a sum of a dissipative drag term $-\gamma U_B(t)$ and a zero-mean randomly fluctuating term $F_B(t)$. As shown in [9], this assumption leads to a mathematical description of $U_B(t)$ as a continuous Markov process in a way that very closely parallels our present treatment of Johnson noise. But a more 'microscopic' approach to Brownian motion would attempt to directly assess the effects of the randomly occurring elastic collisions of the Brownian particle with the molecules of its surrounding fluid. Such an analysis is carried out in [11] for a Brownian particle immersed in an idealized 'one-dimensional gas'. Under the assumption that the gas molecules have a Maxwell–Boltzmann velocity distribution, kinetic theory arguments together with collisional energy–momentum conservation laws are used to derive a rigorous expression for the *probability* that the Brownian particle, moving with speed ν, will collide in the next $\mathrm{d}t$ with some gas molecule in such a way that the particle's speed is instantaneously changed by an amount between ξ and $\xi + \mathrm{d}\xi$. This analysis reveals that $U_B(t)$ is in fact a *jump*-type Markov process: it jumps from one velocity value to

† An equivalent statement of the fluctuation–dissipation theorem, one that is often encountered in the literature, follows upon multiplying equation (26) by itself with t replaced by $t + t'$, averaging the result with the help of equation (8*b*) and then integrating over all t': $R = (2kT)^{-1}\int_{-\infty}^{\infty}\langle V_J(t)V_J(t+t')\rangle\,\mathrm{d}t'$.

another at each discrete collision of the Brownian particle with a gas molecule. But further detailed calculations show that, when we go to observe this jump Markov process $U_B(t)$ on a *macroscopic* time-length scale, $U_B(t)$ *appears* to be a *continuous* Markov process, indeed one of the Ornstein–Uhlenbeck type. The drag coefficient γ, which the direct Langevin analysis regarded as a given parameter, is however now rendered as an *explicit function* of the mass and concentration of the gas molecules and the size of the Brownian particle. So, in view of the strong mathematical connection between Brownian motion and Johnson noise, it seems likely to this writer that any valid microscopic electronic collision model for Johnson noise will simply supplement our key results (24), (25) and (26) by providing an explicit formula for the resistance parameter R in terms of the model's microphysical parameters. Such a formula can of course provide valuable new insights, so efforts to microscopically model Johnson noise are to be strongly encouraged; moreover, such efforts may even lead to some surprises.

4. Shot noise

Turning now to shot noise, we begin by noting that its traditional physical picture is predicated on electrons being emitted by a hot metal. Such emissions are generally observed to occur according to the following rule, where α is some positive constant:

$$\alpha\,\mathrm{d}t = \text{the probability that an electron will be emitted from the metal in the next infinitesimal time increment } \mathrm{d}t. \tag{27}$$

The simplest theoretical basis for this dynamics is the Richardson–Dushman theory of thermionic emission, wherein the electrons inside the metal are treated as an ideal gas, and emissions occur whenever an electron strikes the boundary of the metal with enough energy to escape. As is recounted in standard textbooks [12], α is found to have the form

$$\alpha = CT^{\gamma}\exp(-E_p/kT) \tag{28}$$

where E_p is the photoelectric threshold energy of the metal, k is the Boltzmann's constant, C is a composite of various other physical constants, and γ is either 2 or 1/2 according to whether one assumes the electron gas in the metal obeys Fermi–Dirac statistics or Maxwell–Boltzmann statistics. But a precise formula for α is of less importance to us here than the fact that equation (27), for *some* value of α, accurately describes the way in which electrons are emitted in time.

The emitted electrons are assumed to form a 'parade' which, because of the electron's charge e, constitutes an electrical current. But this electrical current is 'spikey', since the charge flows past any observation point in discrete chunks e; moreover, the current is *stochastic*, since the randomness implicit in the emission law (27) means that we can never predict for sure when a chunk of charge will pass.

To describe this state of affairs mathematically, we begin by defining the integer random variable

$$Z(t_1, t_2) \equiv \text{the number of emitted electrons passing some fixed observation point between times } t_1 \text{ and } t_2 > t_1. \tag{29}$$

Let us denote the probability density function for this random variable by

$$P(z; t_1, t_2) \equiv \mathrm{Prob}\{Z(t_1, t_2) = z\} \qquad (z = 0, 1, \ldots). \tag{30}$$

A formula for P can easily be derived from equation (27) and the laws of probability. For $z = 0$ we have

$$P(0; t_1, t_2 + \mathrm{d}t_2) = P(0; t_1, t_2)(1 - \alpha\,\mathrm{d}t_2).$$

This implies a simple differential equation whose solution is $P(0; t_1, t_2) = \exp[-\alpha(t_2 - t_1)]$. For $z > 0$ we have

$$P(z; t_1, t_2) = \int_{t=t_1}^{t_2} P(z-1; t_1, t)\alpha \, \mathrm{d}t \; P(0; t, t_2).$$

Using the $z = 0$ result, this integral recursion relation may be iterated to obtain the general result

$$P(z; t_1, t_2) = \frac{[\alpha(t_2 - t_1)]^z \exp(-\alpha(t_2 - t_1))}{z!} \qquad (z = 0, 1, \ldots). \tag{31}$$

This establishes $Z(t_1, t_2)$ as a *Poisson* random variable with mean and variance $\alpha(t_2 - t_1)$.

The well known result (31) is a direct, exact consequence of the assumed electron emission dynamics (27), and in a sense it constitutes a complete characterization of the resulting charge transport: it gives the probability that z electrons, and hence a charge ez, will pass a given point in any specified time interval (t_1, t_2). If we naively think of 'electrical current' as simply the amount of charge passing in a unit time, then the fact that the mean and standard deviation of $Z(t_1, t_1 + 1)$ are α and $\sqrt{\alpha}$, respectively, would imply that the mean of the current is $e\alpha$ and the standard deviation of the current is $e\sqrt{\alpha}$. But a *complete* specification of the current here requires that we actually introduce a *dynamical variable* that represents it at each instant of time. It is this necessary but rather subtle step that will bring us into contact with the theory of continuous Markov processes.

To define the *shot noise current* $I_S(t)$ requires that we restrict our observations of the electron parade to a *macroscopic time scale*, for which there exists a time duration δt that, on the one hand, can be regarded as being 'infinitesimally small', but on the other hand is large enough that the average number of electrons passing any observation point in time δt is much larger than 1:

$$\langle Z(t, t + \delta t)\rangle = \alpha \delta t \gg 1. \tag{32}$$

In this case, and *only* in this case, will it make sense to define the 'electrical current at time t' as the charge passing in time $(t, t + \delta t)$ divided by δt:

$$I_S(t) \equiv \frac{eZ(t, t + \delta t)}{\delta t}. \tag{33}$$

Again, δt here is presumed to be *small* enough that this ratio can be regarded as a 'time derivative', yet *large* enough that condition (32) holds.

Now we proceed to quantify the noisy behaviour of this shot noise current $I_S(t)$. We know from equation (31) that $Z(t, t + \delta t)$ is a Poisson random variable with mean and variance $\alpha\delta t$. A well established result in random variable theory says that a *Poisson* random variable with a *very large* mean and variance can be well approximated by a *normal* random variable with the *same* mean and variance. So condition (32) allows us to approximate

$$Z(t, t + \delta t) = \mathcal{N}(\alpha\delta t, \alpha\delta t). \tag{34}$$

Substituting this into the definition (33), and then making use of the normal random variable property (5), we obtain

$$I_S(t) = \frac{e}{\delta t}\mathcal{N}(\alpha\delta t, \alpha\delta t) = \frac{e}{\delta t}[\alpha\delta t + \mathcal{N}(0, \alpha\delta t)] = e\alpha + e\sqrt{\alpha}(\delta t)^{-1}\mathcal{N}(0, \delta t)$$

$$I_S(t) = e\alpha + e\sqrt{\alpha}\mathcal{N}(0, 1/\delta t). \tag{35}$$

But since from our *macroscopic* point of view δt is 'vanishingly small', then by the definition (6) we can replace $\mathcal{N}(0, 1/\delta t)$ in equation (35) with Gaussian white noise:

$$I_S(t) = e\alpha + e\sqrt{\alpha}\Gamma(t). \tag{36}$$

Equation (36) is our final result for the macroscale shot noise current. The critically important fact that $I_S(t)$ does *not* depend on δt, even though the definition (33) of $I_S(t)$ does, is ultimately traceable to the restriction that δt be *macroscopically* infinitesimal yet *microscopically* large enough that condition (32) holds. Since $\langle \Gamma(t) \rangle = 0$, it follows from equation (36) that the mean of the shot noise current is

$$\langle I_S(t) \rangle = e\alpha \equiv \bar{I}_S \tag{37}$$

just as we anticipated earlier in the paragraph following equation (31). But now we are in a position to calculate the *spectral* properties of the shot noise current. Specifically, we seek the current's spectral density function, $S_{I_S}(\nu)$.

To that end, we first calculate the auto-covariance function of $I_S(t)$ (cf equation (12)):

$$\langle I_S(t) I_S(t+t') \rangle = \langle [e\alpha + e\sqrt{\alpha}\Gamma(t)][e\alpha + e\sqrt{\alpha}\Gamma(t+t')] \rangle.$$

In the light of equations (8), this is simply

$$\langle I_S(t) I_S(t+t') \rangle = e^2\alpha^2 + e^2\alpha\delta(t'). \tag{38}$$

This quantity is seen to be independent of t, so the Wiener–Khintchine theorem allows us to compute the spectral density function of $I_S(t)$ as (cf equation (13))

$$S_{I_S}(\nu) = 4\int_0^\infty \langle I_S(t) I_S(t+t') \rangle \cos(2\pi\nu t')\, \mathrm{d}t'.$$

Substituting from equation (38) and performing the easy integration over t', we obtain

$$S_{I_S}(\nu) = 2e^2\alpha^2\delta(\nu) + 2e^2\alpha \qquad (\nu \geqslant 0) \tag{39a}$$

or, using equation (37),

$$S_{I_S}(\nu) = 2\bar{I}_S^2\delta(\nu) + 2e\bar{I}_S \qquad (\nu \geqslant 0). \tag{39b}$$

The first term on the right side of equations (39) displays the zero-frequency or 'DC' part of $S_{I_S}(\nu)$ that comes from the constant term in equation (36). Of greater interest, though, is the second term in equations (39), called the *Schottky term*; it describes a flat or frequency-independent component, which arises from the fluctuating term in equation (36).

If we denote by $Q_S(t)$ the total charge transported by the shot noise current $I_S(t)$ from time 0 to time t, then the fact that these respective variables are related as integral and derivative allows us to deduce from equation (36) that

$$\frac{\mathrm{d}Q_S(t)}{\mathrm{d}t} = e\alpha + e\sqrt{\alpha}\Gamma(t). \tag{40}$$

Comparing this equation with equations (7) and (10*a*), we see that it is the Langevin equation for a *Wiener* process with $a = e\alpha$ and $c = e^2\alpha$. So by equation (10*b*), taking $Q_S(0) = 0$, we have for any $t > 0$,

$$Q_S(t) = \mathcal{N}(e\alpha t, e^2\alpha t). \tag{41}$$

But notice that $I_S(t)$, being the *derivative* of a continuous Markov process, is therefore *not itself* a continuous Markov process. The discontinuous nature of $I_S(t)$ is of course already clear from equation (36), where the Gaussian white noise term precludes continuity.

5. Conclusions

The analysis in sections 3 and 4 reveals some pronounced differences between the Johnson noise current $I_J(t)$ and the shot noise current $I_S(t)$.

First, $I_J(t)$ is a continuous Markov process, specifically an Ornstein–Uhlenbeck process (cf equation (24)), whereas $I_S(t)$ is *not* a continuous Markov process (cf equation (36)). By the same token, the *integral* of $I_S(t)$ is a continuous Markov process, specifically a Wiener process (cf equation (41)), whereas the integral of $I_J(t)$, like the integral of *any* continuous Markov process, in *non*-Markovian.

Second, the spectral density functions of the Johnson noise current in equation (25) and the shot noise current in equation (39) appear to be quite different. Both are arguably flat for moderately low frequencies, but the spectral density function of Johnson noise current has a distinctive $1/f^2$ roll-off at high frequencies, while the spectral density function of the shot noise current has a distinctive delta function spike at zero frequency.

Third, a prominent feature of Johnson noise is its *dissipative* character. The dissipative emf $-RI_J(t)$ saps the Johnson noise current, eventually causing it to relax to a zero-mean (though continually fluctuating) process. And the Johnson noise fluctuation–dissipation relation (26) shows that the dissipative constant R is linked in a very fundamental way to the fluctuations. In contrast to this, the shot noise current $I_S(t)$, although it certainly fluctuates, exhibits *no* dissipative effects; its supporting equations give rise to *no* fluctuation–dissipation relation analogous to the Johnson noise relation (26). So $I_J(t)$ seems to be an intrinsically dissipative process, whereas $I_S(t)$ does not.

The two currents $I_J(t)$ and $I_S(t)$ become a little more alike if we go to the $L \to 0$ limit for the former. In that case, equation (17a) gives $I_J(t) = R^{-1}V_J(t)$, which with equation (26) is

$$I_J(t) = \left(\frac{2kT}{R}\right)^{1/2}\Gamma(t) \qquad (L = 0). \tag{42}$$

This, by equation (14), would give a spectral density function

$$S_{I_J}(\nu) = \frac{4kT}{R} \qquad (L = 0). \tag{43}$$

This $L \to 0$ limit arguably might be realized if the system is viewed on a time scale where the relaxation time L/R is judged to be 'infinitesimally small'. But although both $I_J(t)$ and $I_S(t)$ would then be linearly related to Gaussian white noise, there are still differences: the $I_J(t)$ formula (42) lacks the always present *constant* term in the $I_S(t)$ formula (36); moreover, the white-noise coefficient $(2kT/R)^{1/2}$ in equation (42) seems to be irreconcilably different from the white-noise coefficient $e\sqrt{\alpha}$ in equation (36), at least if the form of α is anything like that in equation (28).

The *tentative* conclusion seems to be that Johnson noise and shot noise are different phenomena. But, as stated in the introduction, this conclusion might be altered if either our definitions of those noises, or the analytical arguments that we made on the basis of those definitions, are amended. Suggestions for and discussions of such amendments could prove enlightening.

References

[1] Johnson J B 1928 *Phys. Rev.* **32** 97
[2] Schottky W 1918 *Ann. Phys., Lpz.* **57** 541
[3] Nyquist H 1928 *Phys. Rev.* **32** 110
[4] Rice S O 1944 *Bell Syst. Tech. J.* **23** 282
Rice S O 1945 *Bell Syst. Tech. J.* **24** 46
[5] de Jong M J M and Beenakker C W J 1997 *Mesoscopic Electron Transport (NATO ASI Series E 345)* ed L Sohn *et al* (Dordrecht: Kluwer) p 225
[6] Blanter Ya M and Buettiker M *Phys. Rep.* at press
[7] Green F and Das M P 1999 *Proc. 2nd Int. Conf. on Unsolved Problems of Noise* ed D Abbott *et al* (New York: AIP) at press

[8] Green F and Das M P *Preprint* cond-mat/9910183
[9] Gillespie D T 1996 *Am. J. Phys.* **64** 225
[10] Thornber K K 1975 *J. Appl. Phys.* **46** 2781
[11] Gillespie D T 1992 *Markov Processes: an Introduction for Physical Scientists* (San Diego: Academic) pp 267–83, 301–12
[12] Leighton R B 1959 *Principles of Modern Physics* (New York: McGraw-Hill) pp 355–7