5 Noise

5.1 Shot noise

Without getting too much into the underlying device physics, shot noise refers to random current fluctuations in electronic devices due to discreteness of charge carriers. The first analysis of shot noise was published by Walter Schottky in 1918 in the context of vacuum tubes, although shot noise also manifests itself in semiconductor devices, such as diodes and transistors, and in optoelectronic devices as a consequence of the fact that electromagnetic radiation is carried by photons.

A vacuum tube is a device that has a pair of terminals or electrodes enclosed in an evacuated glass container. It is used to control the flow of current in a circuit. For example, the simplest type of a vacuum tube, the diode, allows the current to flow in one direction only. These days, there are other types of diodes based on semiconductors, but historically vacuum tubes have been the first active electronic devices. In a diode, electrons are emitted at the cathode and collected at the anode. The duration of each emission event is assumed to be so short that it can be approximated by a unit impulse, and emission events in nonoverlapping time intervals are assumed to be statistically independent. Thus the emission times can be thought of as arrival times of a Poisson process with rate λ, where λ is the average number of electron emissions per unit time. Thus, the current through the diode has a dc component equal to $-q\lambda$, where $q = -1.602 \times 10^{-19}$ Coulombs is the charge of the electron. Shot noise describes the fluctuations around this dc component. The transit of each electron from the cathode to the anode induces a time-varying current, whose shape depends on the device characteristics. Thus, the current through the diode at time $t \ge 0$ is described by the stochastic signal $I = (I_t)_{t \ge 0}$ with

$$I_t = q \sum_{k=1}^{\infty} h(t - T_k), \tag{5.1}$$

where $T = (T_k)_{k \in \mathbb{N}}$ are the arrival times of a Poisson process with rate λ, and $qh(t)$ is the current induced in the diode by an electron emitted at time $t = 0$. We assume that $h(t) = 0$ for $t < 0$ (no current before an emission) and that $\int_{-\infty}^{\infty} h(t)\mathrm{d}t = \int_0^{\infty} h(t)\mathrm{d}t = 1$ (conservation of charge).

The mean and the variance of I_t are given by Campbell's theorem:

$$m_I(t) = \mathbf{E}[I_t] = q\lambda \int_{-\infty}^{\infty} h(\tau)\mathrm{d}\tau = q\lambda \int_0^{\infty} h(\tau)\mathrm{d}\tau = q\lambda$$

(which is the dc component) and

$$\sigma_I^2(t) = \mathrm{Var}[I_t] = q^2\lambda \int_{-\infty}^{\infty} h^2(\tau)\mathrm{d}\tau = q^2\lambda \int_0^{\infty} h^2(\tau)\mathrm{d}\tau$$

(which measures the fluctuations due to randomly timed emissions). Denoting by H the Fourier transform of h, we note that

$$\int_0^{\infty} h(\tau)\mathrm{d}\tau = H(0)$$

and, by Parseval's theorem,

$$\int_0^\infty h^2(\tau)\mathrm{d}\tau = \frac{1}{2\pi}\int_{-\infty}^\infty |H(\omega)|^2 \mathrm{d}\omega.$$

Therefore,

$$m_I(t) = q\lambda H(0) \quad \text{and} \quad \sigma_I^2(t) = \frac{q^2\lambda}{2\pi}\int_{-\infty}^\infty |H(\omega)|^2 \mathrm{d}\omega.$$

However, in order to get the full picture of shot-noise current, we will need to compute the autocorrelation of I and its power spectral density.

In order to compute $R_I(t, t+\tau)$, we will use the same approach that we relied on to prove Campbell's formula for the variance of I_t. Namely, we first write I as the output of an LTI system with impulse response h and with the input $Z = (Z_t)_{t\geq 0}$ given by

$$Z_t = q\sum_{k=1}^\infty \delta(t - T_k).$$

Here, Z is a train of unit impulses originating at the arrival times $T_1, T_2, \ldots$. Each impulse corresponds to the emission of a single electron at the cathode. We have already computed the autocorrelation of Z when proving Campbell's theorem: it is given by

$$R_Z(\tau) = q^2\left(\lambda^2 + \lambda\delta(\tau)\right). \tag{5.2}$$

Therefore,

$$\begin{aligned}
R_I(\tau) &= \tilde{h} * h * R_Z(\tau) \\
&= \int_{-\infty}^\infty \int_{-\infty}^\infty \tilde{h}(\tau - s)h(s-t)R_Z(t)\,\mathrm{d}s\,\mathrm{d}t \\
&= q^2\lambda^2 \int_{-\infty}^\infty \int_{-\infty}^\infty \tilde{h}(\tau - s)h(s-t)\,\mathrm{d}s\,\mathrm{d}t + q^2\lambda \int_{-\infty}^\infty \int_{-\infty}^\infty \tilde{h}(\tau - s)h(s-t)\delta(t)\,\mathrm{d}s\,\mathrm{d}t \\
&= q^2\lambda\left(\int_{-\infty}^\infty h(t)\mathrm{d}t\right)^2 + q^2\lambda\tilde{h} * h(\tau) \\
&= q^2\lambda + q^2\lambda\tilde{h} * h(\tau),
\end{aligned}$$

where we have used the definition of the time reversal $\tilde{h}(t) = h(-t)$, as well as the fact that $\int_{-\infty}^\infty h(t)\mathrm{d}t = 1$. The power spectral density of I is the Fourier transform of R_I:

$$S_I(\omega) = 2\pi q^2\lambda\delta(\omega) + q^2\lambda|H(\omega)|^2. \tag{5.3}$$

Note the ttwo terms in (5.3): the first term is the dc term, while the second term is due to the filtering effect of h on Z.

It is useful to consider an example: Assuming that each electron moves from the cathode to the anode with constant velocity and that a constant time t_0 elapses between the emission and the collection of each electron, it can be shown[1] that the impulse response h has the form

$$h(t) = \begin{cases} \frac{1}{t_0}, & 0 \le t \le t_0 \\ 0, & \text{otherwise} \end{cases}.$$

Taking the Fourier transform, we get

$$\begin{aligned} H(\omega) &= \int_{-\infty}^{\infty} h(t) e^{-\mathrm{i}\omega t} \mathrm{d}t \\ &= \frac{1}{t_0} \int_0^{t_0} e^{-\mathrm{i}\omega t} \mathrm{d}t \\ &= \frac{-1}{\mathrm{i}\omega t_0} e^{-\mathrm{i}\omega t} \Big|_0^{t_0} \\ &= \frac{1 - e^{-\mathrm{i}\omega t_0}}{\mathrm{i}\omega t_0} \\ &= \frac{2e^{-\mathrm{i}\omega t_0/2}}{\omega t_0} \frac{e^{\mathrm{i}\omega t_0/2} - e^{-\mathrm{i}\omega_0 t/2}}{2\mathrm{i}} \\ &= e^{-\mathrm{i}\omega t_0/2} \frac{\sin\left(\frac{\omega t_0}{2}\right)}{\frac{\omega t_0}{2}} \\ &= e^{-\mathrm{i}\omega t_0/2} \operatorname{sinc}\left(\frac{\omega t_0}{2\pi}\right). \end{aligned}$$

Substituting this into (5.3), we get

$$S_I(\omega) = 2\pi q^2 \lambda \delta(\omega) + q^2 \lambda \operatorname{sinc}^2\left(\frac{\omega t_0}{2\pi}\right).$$

5.2 Johnson–Nyquist noise

Shot noise depends only on the current, and is due entirely to the discreteness of charge carriers. Another source of noise in devices is due to thermal agitation of charge carriers, and its effect increases with resistance and with temperature. Let us first consider the case of a noisy resistor with resistance R at absolute temperature T. Using an argument based on the so-called equipartition theorem from statistical physics, Nyquist showed that such a noisy resistor can be described by a Thévenin equivalent model consisting of a *noiseless* resistor with resistance R in series with a *noisy* voltage source whose voltage is given by a Gaussian stochastic signal $E = (E_t)_{t\in\mathbb{R}}$ with zero mean and the white-noise autocorrelation function $R_E(\tau) = 2kTR\delta(\tau)$. Here, T is the temperature in Kelvin (e.g., 290 K corresponds to room temperature), and $k = 1.3806 \times 10^{-23}$ J/K is the Boltzmann constant.

[1] See, e.g., Section 4.1 in Hermann Haus, *Electromagnetic Noise and Quantum Optical Measurements*, Springer, Berlin, 2000.

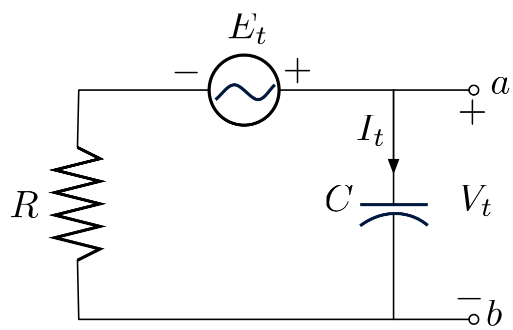

Figure 1: An RC circuit with a noisy resistor.

Alternatively, we can use the Norton equivalent, where the noiseless resistor R is connected in parallel with a noisy current source $J = (J_t)_{t\in\mathbb{R}}$ given by $J_t = \frac{E_t}{R}$, so that $R_J(\tau) = 2kTG\delta(\tau)$, where $G = 1/R$ is the conductance. Consequently, the Nyquist voltage source has the power spectral density $S_E(\omega) = 2kTR$, while the Nyquist current source has $S_J(\omega) = 2kTG$. Of course, ideal white noise is unphysical (recall that it has infnite average power), and it is often sufficient to assume that the noisy voltage (or current) power spectral density is constant and nonzero only over some finite bandwidth. In that case, we will have

$$S_E(\omega) = \begin{cases} 2kTR, & |\omega| \le 2\pi B \\ 0, & \text{otherwise} \end{cases} \tag{5.4}$$

where B is the bandwidth in Hz, and so

$$\begin{aligned} \mathbf{E}[E_t^2] &= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_E(\omega)\mathrm{d}\omega \\ &= \frac{1}{2\pi}\int_{-2\pi B}^{2\pi B} 2kTR\,\mathrm{d}\omega \\ &= 4kTRB. \end{aligned}$$

Nyquist obtained this result by modeling the resistor as a transmission line and by considering the thermodynamics of the Brownian motion of electrons in the resistor.[2] Here, we will take it for granted that E is a Gaussian white-noise stochastic signal with power spectral density $S_E(\omega) = S_0$ and derive the constant S_0 using a bit of circuit analysis, input-output relations for LTI systems with stochastic inputs, and a result from thermodynamics known as the *equipartition theorem*.

Let us think about how we would go about measuring the voltage fluctuations due to thermal noise. Consider the RC circuit shown in Figure 1. Since E_t is a noisy voltage source, all currents and voltages in this circuit are stochastic signals. We connect a capacitor in parallel with the noisy resistor, and will measure the voltage across the terminals a and b, i.e., the voltage across the capacitor. Kicrhhoff's voltage law gives

$$-I_tR + E_t - V_t = 0, \tag{5.5}$$

[2] H. Nyquist, "Thermal agitation of electric charge in conductors," *Physical Review*, vol. 32, pp. 110–113, 1928

where the current through the capacitor is given by $I_t = C\frac{\mathrm{d}V_t}{\mathrm{d}t}$. Consequently,

$$RC\frac{\mathrm{d}V_t}{\mathrm{d}t} + V_t = E_t,$$

which means that the transfer function from E to V is given by

$$H(\omega) = \frac{1}{1+\mathrm{i}\omega RC}.$$

Thus, the power spectral density of V is given by

$$\begin{aligned} S_V(\omega) &= |H(\omega)|^2 S_E(\omega) \\ &= \frac{S_E(\omega)}{1+\omega^2(RC)^2} \\ &= \frac{S_0}{1+\omega^2(RC)^2}. \end{aligned} \tag{5.6}$$

From this, we can compute the average energy stored in the capacitor:

$$\begin{aligned} \frac{1}{2}C\mathbf{E}[V_t^2] &= \frac{1}{2}C\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty} S_V(\omega)\mathrm{d}\omega \\ &= \frac{CS_0}{4\pi}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{1+\omega^2(RC)^2}. \end{aligned} \tag{5.7}$$

The integral can be computed in closed form: consider the trigonometric substitution $x = \tan\theta$. Then $\mathrm{d}x = (1/\cos^2\theta)\mathrm{d}\theta$, and

$$\begin{aligned} \int\frac{\mathrm{d}x}{1+x^2} &= \int\frac{1}{1+\tan^2\theta}\frac{\mathrm{d}\theta}{\cos^2\theta} \\ &= \int\frac{\mathrm{d}\theta}{\sin^2\theta+\cos^2\theta} \\ &= \int\mathrm{d}\theta \\ &= \theta + \mathrm{const} \\ &= \arctan x + \mathrm{const}. \end{aligned}$$

Therefore,

$$\int_{-\infty}^{\infty}\frac{\mathrm{d}x}{1+x^2} = \int_{\pi/2}^{\pi/2}\mathrm{d}\theta = \pi.$$

Using this in Eq. (5.7), we obtain

$$\frac{1}{2}C\mathbf{E}[V_t^2] = \frac{S_0}{4R}. \tag{5.8}$$

Now we invoke the equipartition theorem, which says that, in a system consisting of many independent degrees of freedom in thermal equilibrium at temperature T, the average energy stored in each degree of freedom is equal to $\frac{1}{2}kT$, where k is the Boltzmann constant. Treating the RC circuit as a single degree of freedom and using (5.8), we get

$$\frac{1}{2}C\,\mathbb{E}[V_t^2] = \frac{S_0}{4R} = \frac{1}{2}kT,$$

which gives $S_0 = 2kTR$.

Another way to express (5.6), with $S_0 = 2kTR$, is as follows: Let $Z(\omega)$ denote the impedance of the circuit between the terminals a and b. Viewed from these terminals, the resistor R and the capacitor C are connected in parallel, and therefore

$$\begin{aligned}\frac{1}{Z(\omega)} &= \frac{1}{Z_R(\omega)} + \frac{1}{Z_C(\omega)} \\ &= \frac{1}{R} + \mathrm{i}\omega C,\end{aligned}$$

or, equivalently,

$$Z(\omega) = \frac{R}{1+\mathrm{i}\omega RC}.$$

Then, for the ideal Johnson–Nyquist noise source, we have

$$S_V(\omega) = 2kT\,\mathrm{Re}\,Z(\omega). \tag{5.9}$$

We will now show that Eq. (5.9) is a special case of a more general theorem of Nyquist: The voltage across any pair of terminals a,b in a network consisting of noisy resistors at temperature T, inductors, and capacitors has the power spectral density

$$S_V(\omega) = 2kT\,\mathrm{Re}\,Z(\omega) \tag{5.10}$$

for ideal Johnson–Nyquist noise, where $Z(\omega)$ is the impedance of the network across the termimal pair a,b.

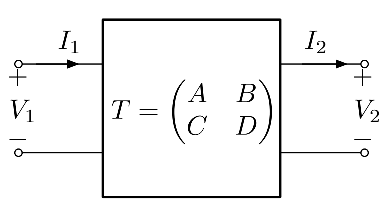

Figure 2: A two-port circuit.

In order to prove Nyquist's theorem, we first need to introduce some background concepts. A *two-port network*, shown in Figure 2, is a circuit with two ports, i.e., two pairs of terminals that can be connected to external circuits. Each port has a voltage and a current, and the relationship between the voltages and the currents is specified by the so-called *transfer matrix*

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix},$$

where V_1 is the voltage across port 1, I_1 is the current into port 1, V_2 is the voltage across port 2, and I_2 is the current out of port 2. It is important to keep in mind that this is a *frequency-domain* description. We say that the two-port circuit is:

- *reciprocal* if the transfer function from I_1 to V_2 is equal to the transfer function from $-I_2$ to V_1:

$$\frac{V_2(\omega)}{I_1(\omega)} = -\frac{V_1(\omega)}{I_2(\omega)}, \qquad \forall \omega$$

  (the minus sign is there because I_2 is the current flowing out of port 2);

- *lossless* if the power supplied at port 1 is equal to the power that exits at port 2, and vice versa.

A network that consists only of resistors, inductors, and capacitors is reciprocal; a network that consists only of inductors and capacitors is lossless. We will assume, for simplicity, that our network has only one resistor, and all other elements are lossless. If we represent the noisy resistor by its Norton equivalent consisting of a noiseless resistor in parallel with a noisy current source $J = (J_t)_{t \in \mathbb{R}}$, then we can view J as a (noisy) current $I_1 = (I_{1,t})$ into port 1 of a two-port network and the voltage across the resistor as the noisy voltage $V_1 = (V_{1,t})$ across port 1, and where we associate port 2 with the terminals a and b — see Figure 3. Thus, the voltage signal of interest is the voltage $V_2 = (V_{2,t})$ across port 2. The power flowing into port 2 is equal to $|I_2(\omega)|^2 \operatorname{Re} Z(\omega)$, and, since the

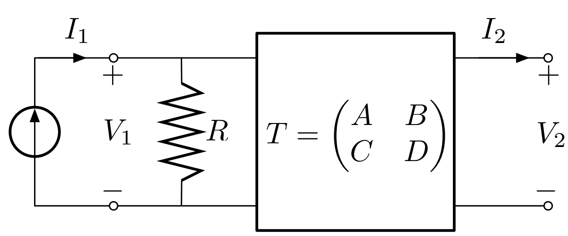

Figure 3: A general passive network with a noisy resistor; the T-matrix represents the rest of the network, which consists only of inductors and capacitors.

network connected to the resistor is lossless by assumption, all of this power is dissipated at the resistor. That is,

$$|I_2(\omega)|^2 \operatorname{Re} Z(\omega) = \frac{|V_1(\omega)|^2}{R}. \tag{5.11}$$

Let H denote the transfer function from I_1 to V_2. By reciprocity, it is equal to the transfer function from $-I_2$ to V_1. Using this fact together with (5.11), we can write

$$|H(\omega)|^2 = \left|\frac{V_2(\omega)}{I_1(\omega)}\right|^2 = \left|\frac{V_1(\omega)}{I_2(\omega)}\right|^2 = R\,\mathrm{Re}\,Z(\omega). \tag{5.12}$$

Since the power spectral density of the voltage $V = V_2$ is given by

$$S_V(\omega) = |H(\omega)|^2 S_J(\omega),$$

we can use (5.12) and the fact that $S_J(\omega) = \frac{S_E(\omega)}{R^2}$ to obtain Eq. (5.10).

5.3 Bonus section: amplifier noise

A good illustration of the theory of noise sources in circuits is the analysis of *amplifier noise*. An amplifier is an active device that provides a power gain relative to its input. A noisy amplifier, just like its name suggests, also introduces additional noise at the output. In order to discuss amplifiers, we first must introduce the notions of a noisy source and exchangeable power. A source, shown in Figure 4 (left), is a one-port circuit (i.e., a circuit with a single pair of terminals) that has the voltage-current characteristic

$$V = Z_{\mathrm{S}} I + E, \tag{5.13}$$

where V is the voltage across the source's terminals, Z_{S} is the source impedance, I is the current flowing into the source, and E is the open-circuit voltage of the source. We say that the source is noisy if its open-circuit voltage E is a stochastic signal. For example, if the source is a noisy resistor with resistance R at temperature T, then we can take E to be the bandlimited Johnson–Nyquist voltage noise.

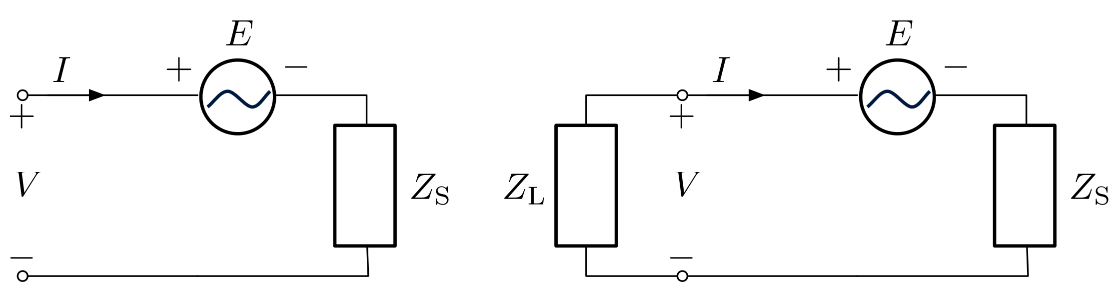

Figure 4: The equivalent circuit of a source (left) and a source connected to a load (right).

Suppose that we connect the source to a load with load impedance Z_{L}, see Figure 4 (right). The average power delivered to the load at time t is given by

$$P_{\mathrm{L},t} = \mathbf{E}[|I_t|^2]\,\mathrm{Re}\,Z_{\mathrm{L}},$$

where the Fourier transform of I_t can be computed from the Kirchhoff's voltage law:

$$I(\omega) = -\frac{E(\omega)}{Z_{\mathrm{L}} + Z_{\mathrm{S}}}.$$

Therefore,

$$P_{\mathrm{L},t} = \frac{\mathbf{E}[|E_t|^2]\,\mathrm{Re}\,Z_{\mathrm{L}}}{|Z_{\mathrm{S}} + Z_{\mathrm{L}}|^2}. \tag{5.14}$$

If E is weakly stationary, then $\mathbf{E}[|E_t^2|]$ does not depend on t, and we can express (5.14) as

$$P_{\mathrm{L}} = \frac{R_E(0)\,\mathrm{Re}\,Z_{\mathrm{L}}}{|Z_{\mathrm{S}} + Z_{\mathrm{L}}|^2}, \tag{5.15}$$

where $R_E(\tau)$ denotes the autocorrelation function of E. The quantity in (5.15) is a function of the load impedance Z_{L}. *Exchangeable power*, denoted by P_{ex}, is the global *extremum* (maximum or minimum) value of the power fed to the load by the source. Exchangeable power can be positive (the source dissipating power into the load) or negative (the load dissipating power into the source). To compute it, we first separate the load and the source impedances into the resistive and the reactive components: $Z_{\mathrm{L}} = R_{\mathrm{L}} + \mathrm{i}X_{\mathrm{L}}$ and $Z_{\mathrm{S}} = R_{\mathrm{S}} + \mathrm{i}X_{\mathrm{S}}$. Then

$$P_{\mathrm{L}} = \frac{R_E(0)R_{\mathrm{L}}}{(R_{\mathrm{L}} + R_{\mathrm{S}})^2 + (X_{\mathrm{L}} + X_{\mathrm{S}})^2},$$

and since

$$\frac{\partial P_{\mathrm{L}}}{\partial X_{\mathrm{L}}} = \frac{-2R_E(0)(X_{\mathrm{L}} + X_{\mathrm{S}})}{\left((R_{\mathrm{L}} + R_{\mathrm{S}})^2 + (X_{\mathrm{L}} + X_{\mathrm{S}})^2\right)^2},$$

we see that $X_{\mathrm{L}} = -X_{\mathrm{S}}$ is a necessary condition for extremum. Therefore, we assume that $X_{\mathrm{L}} = -X_{\mathrm{S}}$, and so now we need to compute the extremum of

$$\frac{R_E(0)R_{\mathrm{L}}}{(R_{\mathrm{L}} + R_{\mathrm{S}})^2}$$

over R_{L}. In this case, it is a simple exercise in calculus to show that, upon setting $R_{\mathrm{L}} = R_{\mathrm{S}}$, the power delivered to the load achieves its global maximum if $R_{\mathrm{S}} > 0$ and global minimum if $R_{\mathrm{L}} < 0$. In either case, the global extremum value is the exchangeable power, given by

$$P_{\mathrm{ex}} = \frac{R_E(0)}{2(Z_{\mathrm{S}} + Z_{\mathrm{S}}^*)}, \tag{5.16}$$

and it is achieved by matching the load impedance to the source impedance, i.e., $Z_{\mathrm{L}} = Z_{\mathrm{S}}^*$.

Now suppose that we connect the source to one port of a two-port circuit with transfer matrix T, as shown in Figure 5. The *exchangeable gain* G of the resulting network is defined as the ratio of

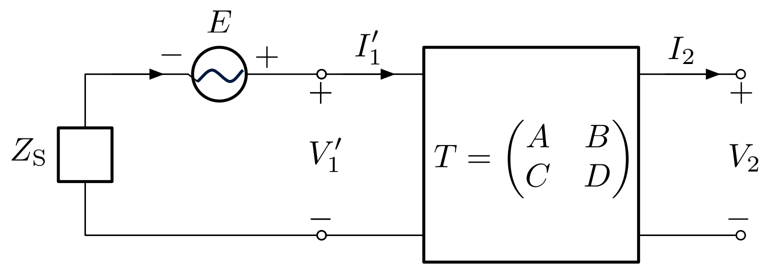

Figure 5: A noisy source connected to a noiseless two-port circuit.

the output exchangeable power $P_{ex,2}$ (i.e, at port 2) to the input exchangeable power $P_{ex,1}$ (i.e., at port 1). The latter is given by (5.16):

$$P_{ex,1} = \frac{R_E(0)}{2(Z_S + Z_S^*)}.$$

To compute $P_{ex,2}$, we need a voltage-current characteristic like (5.13) to relate V_2, I_2, and E. To that end, we first write down the relation between the input and output currents and voltages for our two-port:

$$V_1' = AV_2 + BI_2, \tag{5.17a}$$
$$I_1' = CV_2 + DI_2. \tag{5.17b}$$

Substituting the expressions (5.17) into Kirchhoff's voltage law $V_1' = E - I_1' Z_S$, we obtain

$$AV_2 + BI_2 = -(CV_2 + DI_2)Z_S + E.$$

Rearranging, we get

$$V_2 = -\frac{B + DZ_S}{A + CZ_S} I_2 + \frac{1}{A + CZ_S} E,$$

so, looking from port 2, we can view the whole circuit as a source with effective impedance $\tilde{Z}_S = \frac{B+DZ_S}{A+CZ_S}$ and open-circuit voltage $\tilde{E} = \frac{E}{A+CZ_S}$. Therefore, the output exchangeable power is given by

$$P_{ex,2} = \frac{R_{\tilde{E}}(0)}{2(\tilde{Z}_S + \tilde{Z}_S^*)}. \tag{5.18}$$

Now, since

$$R_{\tilde{E}}(0) = \frac{R_E(0)}{|A + CZ_S|^2}$$

and

$$\tilde{Z}_S + \tilde{Z}_S^* = \frac{B + DZ_S}{A + CZ_S} + \frac{B^* + D^* Z_S^*}{A^* + C^* Z_S^*}$$
$$= \frac{(B + DZ_S)(A + CZ_S)^* + (B + DZ_S)^*(A + CZ_S)}{|A + CZ_S|^2},$$

we get from (5.18)

$$P_{ex,2} = \frac{R_E(0)}{2\left[(B + DZ_S)(A + CZ_S)^* + (B + DZ_S)^*(A + CZ_S)\right]},$$

and can now write down the exchangeable gain:

$$G = \frac{P_{ex,2}}{P_{ex,1}} = \frac{Z_S + Z_S^*}{(B + DZ_S)(A + CZ_S)^* + (B + DZ_S)^*(A + CZ_S)} \tag{5.19}$$

This expression looks rather formidable, but it can be written down more succinctly using matrix notation. Define the following complex vector v and matrix S:

$$v \triangleq \begin{pmatrix} 1 \\ Z_S^* \end{pmatrix}, \qquad S \triangleq 2\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and recall that the *Hermitian conjugate* $A^\dagger$ of a matrix A is obtained by taking the complex conjugate of each entry, followed by taking the transpose. Then we can write

$$G = \frac{v^\dagger S v}{v^\dagger T S T^\dagger v} \tag{5.20}$$

(prove this!). An amplifier has $G \geq 1$.

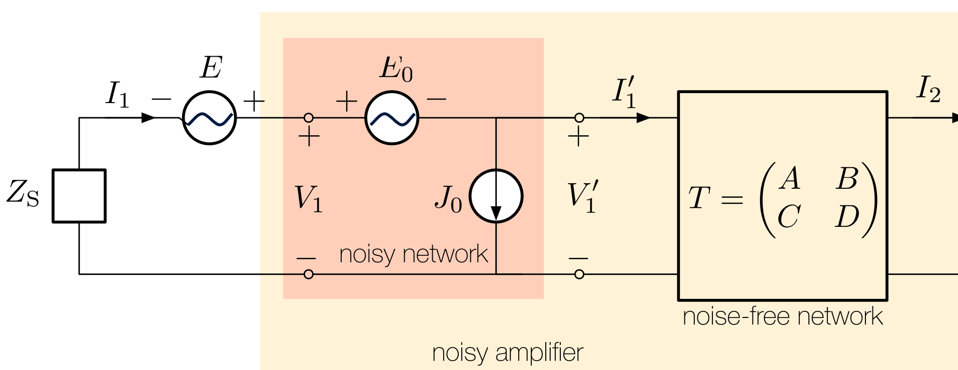

Figure 6: An equivalent circuit model of a noisy amplifier.

Now we are going to consider the effect of internal noise in the amplifier. Amplifier noise is typically modeled by including internal noisy voltage and current sources into the amplifier network, as shown in Figure 6. We assume that the noisy voltage E_0 and the noisy current J_0 are independent of the noisy source voltage E. The effect of these internal noise sources manifests itself at the output of the amplifier. To quantify it, one typically uses the so-called *noise figure* F, which is defined as the ratio of the output signal-to-noise ratio to the input signal-to-noise ratio. In our case, the noise figure is given by the ratio of the variance of the voltage V_1' at the input to the noiseless part of the amplifier and the variance of the source voltage E. To compute the variance of $V_{1,t}'$, we first replace the combination of the source and the noisy amplifier network by its Thévenin equivalent circuit with source impedance Z_S in series with the Thévenin voltage source $E^{\mathrm{Th}} = E - E_0 - JZ_S$. Then $V_1' = E^{\mathrm{Th}}$, and, since E is independent of E_0 and J_0,

$$\mathbf{E}[|V_{1,t}'|^2] = \mathbf{E}[|E_t|^2] + \mathbf{E}[|E_{0,t} + J_{0,t}Z_S|^2].$$

Assuming that all noise sources are jointly weakly stationary, we have $\mathbf{E}[|E_t|^2] = R_E(0)$ and

$$\begin{aligned}\mathbf{E}[|E_{0,t} + J_{0,t}Z_S|^2] &= \mathbf{E}[|E_{0,t}|^2] + 2\mathrm{Re}\,\mathbf{E}[E_{0,t}^* J_{0,t}] + \mathbf{E}[|J_{0,t}|^2] \\ &= R_{E_0}(0) + 2\mathrm{Re}, R_{E_0J_0}(0) + R_{J_0}(0).\end{aligned}$$

Therefore, the amplifier noise figure is given by

$$\begin{aligned}F &= \frac{\mathbf{E}[|V_{1,t}'|^2]}{\mathbf{E}[|E_t|^2]} \\ &= \frac{R_E(0) + R_{E_0}(0) + 2\mathrm{Re}, R_{E_0J_0}(0) + R_{J_0}(0)}{R_E(0)} \\ &= 1 + \frac{R_{E_0}(0) + 2\mathrm{Re}, R_{E_0J_0}(0) + R_{J_0}(0)}{R_E(0)}.\end{aligned}$$

Note that $F \geq 1$. If the input source consists of an impedance Z_S in series with a bandlimited Johnson–Nyquist noise source with bandwidth B at temperature T, then $R_E(0) = 4kTB\mathrm{Re}Z_S$, and therefore we can write down the so-called *excess noise figure*

$$\begin{aligned}F - 1 &= \frac{R_{E_0}(0) + 2\mathrm{Re}, R_{E_0J_0}(0) + R_{J_0}(0)}{4kTB\,\mathrm{Re}\,Z_S} \\ &= \frac{R^{\mathrm{Th}}(0)}{2kTB(Z_S + Z_S^*)},\end{aligned}$$

where R^{Th} is the autocorrelation function of E^{Th}. Thus, $F - 1$ is equal to the exchangeable power $P_{\mathrm{ex}}^{\mathrm{Th}}$ at the source with impedance Z_S and noisy open-circuit voltage E^{Th}, divided by kTB:

$$F - 1 = \frac{P_{\mathrm{ex}}^{\mathrm{Th}}}{kTB}.$$

Now that we have obtained the expressions for the exchangeable gain G and the noise figure F, we can compute the so-called *amplifier noise measure*:

$$
\begin{aligned}
M &\triangleq \frac{F-1}{1-1/G} \\
&= \frac{\frac{P_{\text{ex}}^{\text{Th}}}{kTB}}{\frac{v^\dagger S v}{v^\dagger T S T^\dagger v}}.
\end{aligned}
$$

For high-gain amplifiers ($G \gg 1$), $M \approx F - 1$.

5.4 $1/f$ noise

Both shot noise and the Johnson–Nyquist noise are due to fundamental physical phenomena, namely discreteness of charge carriers and their thermal agitation. These two effects would still be present in ideal devices. However, measurements reveal another noise phenomenon due to the presence of impurities. This phenomenon, referred to as $1/f$ noise for reasons that will become clear shortly, manifests itself as random fluctuations of the conductance of the device due to spontaneous generation and recombination of electrons and holes. These fluctuations are especially noticeable at low frequencies, which is why frequency upconversion is often employed in low-noise electronic devices. In fact, experiments have shown that the power spectrum of these conductance fluctuations behaves as $1/f$, where $f = \omega/2\pi$ is the frequency in Hertz. This is the reason for the name "$1/f$ noise."

As we will now see, a simple model of $1/f$ noise can be derived by starting from a toy Markov-chain model of generation-recombination noise in a nearly pure semiconductor. Consider a continuous-time stochastic signal $X = (X_t)_{t\in\mathbb{R}}$ with binary state space $\mathsf{X} = \{-1, +1\}$. The state -1 corresponds to the spontaneous generation of an electron, the state $+1$ corresponds to the spontaneous generation of a hole, and we will arrange things in such a way that, on average, there are no charge fluctuations, so $\mathbf{P}[X_t = -1] = \mathbf{P}[X_t = +1] = \frac{1}{2}$. For $s \le t$, let $M_{s,t}(-,-)$ denote the conditional probability that the $X_t = -1$ given $X_s = -1$, and define $M_{s,t}(-,+)$, $M_{s,t}(+,-)$, and $M_{s,t}(+,+)$ analogously. We will first construct a time-discretized approximation, and then look at the continuous-time limit. Thus, pick a sufficiently small value $h > 0$ and assume that the state can change at times $\dots, -2h, -h, 0, h, 2h, \dots$, and that the state transitions at different times are independent. Let $p_t(-) \triangleq \mathbf{P}[X_t = -1]$. We will assume that, for small enough h, we can approximate $M_{t,t+h}(-,+) = M_{t,t+h}(+,-) \simeq \alpha h$ for some $\alpha > 0$, i.e., α is the rate of state transitions per unit time. Then, for any $t \in \mathbb{R}$ and any $h < 1/\alpha$, we have

$$
\begin{aligned}
p_{t+h}(-) &= p_t(-)M_{t,t+h}(-,-) + p_t(+)M_{t,t+h}(+,-) \\
&= (1-\alpha h)p_t(-) + \alpha h p_t(+),
\end{aligned}
$$

and an analogous derivation gives

$$
p_{t+h}(+) = \alpha h p_t(-) + (1-\alpha h)p_t(+).
$$

In matrix form, we obtain

$$
\begin{pmatrix} p_{t+h}(-) & p_{t+h}(+) \end{pmatrix} = \begin{pmatrix} p_t(-) & p_t(+) \end{pmatrix} \begin{pmatrix} (1-\alpha)h & \alpha h \\ \alpha h & 1-\alpha h \end{pmatrix}, \tag{5.21}
$$

which, upon rearranging, becomes

$$\frac{1}{h}\begin{pmatrix} p_{t+h}(-)-p_t(-) & p_{t+h}(+)-p_t(+) \end{pmatrix} = \begin{pmatrix} p_t(-) & p_t(+) \end{pmatrix}\begin{pmatrix} -\alpha & \alpha \\ \alpha & -\alpha \end{pmatrix}.$$

Taking the continuous-time limit $h \to 0$, we end up the first-order matrix ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} p_t(-) & p_t(+) \end{pmatrix} = \begin{pmatrix} p_t(-) & p_t(+) \end{pmatrix}\begin{pmatrix} -\alpha & \alpha \\ \alpha & -\alpha \end{pmatrix}. \tag{5.22}$$

It can be solved by diagonalizing the matrix on the right, and, for any $t \in \mathbb{R}$ and any $\tau \ge 0$, the solution is given by

$$\begin{pmatrix} p_{t+\tau}(-) & p_{t+\tau}(+) \end{pmatrix} = \begin{pmatrix} p_t(-) & p_t(+) \end{pmatrix}\begin{pmatrix} \frac{1+e^{-2\alpha\tau}}{2} & \frac{1-e^{-2\alpha\tau}}{2} \\ \frac{1-e^{-2\alpha\tau}}{2} & \frac{1+e^{-2\alpha\tau}}{2} \end{pmatrix}. \tag{5.23}$$

For $s \le t$, define the matrix

$$M_{s,t} = \begin{pmatrix} M_{s,t}(-,-) & M_{s,t}(-,+) \\ M_{s,t}(+,-) & M_{s,t}(+,+) \end{pmatrix} \triangleq \begin{pmatrix} \frac{1+e^{-2\alpha(t-s)}}{2} & \frac{1-e^{-2\alpha(t-s)}}{2} \\ \frac{1-e^{-2\alpha(t-s)}}{2} & \frac{1+e^{-2\alpha(t-s)}}{2} \end{pmatrix}. \tag{5.24}$$

Then $M_{s,s} = I$, $p_t = p_s M_{s,t}$, and a simple calculation shows that the probability vector $\pi = (1/2, 1/2)^T$ is invariant under all $M_{s,t}$, i.e., if $p_s(-) = p_s(+) = 1/2$, then $p_t(-) = p_t(+) = 1/2$ as well. We will assume, therefore, that $p_t(-) = p_t(+) = 1/2$ for all t, corresponding to no excess electrons or holes on average. Since

$$\begin{aligned}
\mathbf{P}[X_t = X_{t+\tau}] &= \mathbf{P}[X_t = -1, X_{t+\tau} = -1] + \mathbf{P}[X_t = +1, X_{t+\tau} = +1] \\
&= p_t(-)M_{t,t+\tau}(-,-) + p_t(+)M_{t,t+\tau}(+,+) \\
&= \frac{1+e^{-2\alpha\tau}}{2}
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{P}[X_t = -X_{t+\tau}] &= \mathbf{P}[X_t = -1, X_{t+\tau} = +1] + \mathbf{P}[X_t = +1, X_{t+\tau} = -1] \\
&= p_t(-)M_{t,t+\tau}(-,+) + p_t(+)M_{t,t+\tau}(+,-) \\
&= \frac{1-e^{-2\alpha\tau}}{2},
\end{aligned}$$

the autocorrelation function of X is given by

$$\begin{aligned}
R_X(t, t+\tau) &= \mathbf{E}[X_t X_{t+\tau}] \\
&= \mathbf{P}[X_t = X_{t+\tau}] - \mathbf{P}[X_t = -X_{t+\tau}] \\
&= e^{-2\alpha\tau}.
\end{aligned}$$

A similar calculation shows that $R_X(t-\tau,t) = e^{-2\alpha\tau}$, so we obtain

$$R_X(\tau) = e^{-2\alpha|\tau|}. \tag{5.25}$$

It is convenient to define the *relaxation time* $t_0 \triangleq 1/2\alpha$ and express (5.25) as $R_X(\tau) = e^{-|\tau|/t_0}$. Now let us compute the power spectral density of X:

$$\begin{aligned}
S_X(\omega) &= \int_{-\infty}^{\infty} R_X(\tau)e^{-\mathrm{i}\omega\tau}\mathrm{d}\tau \\
&= \int_{-\infty}^{\infty} e^{-|\tau|/t_0}e^{-\mathrm{i}\omega\tau}\mathrm{d}\tau \\
&= \int_{-\infty}^{0} e^{(1/t_0-\mathrm{i}\omega)\tau}\mathrm{d}\tau + \int_{0}^{\infty} e^{-(1/t_0+\mathrm{i}\omega)\tau}\mathrm{d}\tau \\
&= \frac{1}{1/t_0-\mathrm{i}\omega} + \frac{1}{1/t_0+\mathrm{i}\omega} \\
&= \frac{2/t_0}{1/t_0^2+\omega^2} \\
&= \frac{2t_0}{1+(t_0\omega)^2}.
\end{aligned}$$

In this context, it is customary to work with the frequency in Hertz, $f = \omega/2\pi$, so we obtain

$$S_X(f) = \frac{2t_0}{1+(2\pi t_0 f)^2}, \tag{5.26}$$

which is known as the *Lorentzian power spectrum*, see Figure 7.

The Lorentzian power spectrum describes the conductance fluctuations in a nearly pure semiconductor, where t_0 is the characteristic timescale of electron-hole recombination. However, in a typical semiconductor device with impurities, there are multiple generation-recombination processes going on, each with its own relaxation time. A good physical model, as it turns out, results if we assume that the relaxation time is a nonnegative *random variable* with a pdf g, such that, for any $t' \le t_0 \le t''$ and any n times $t_1 < t_2 < \ldots < t_n$,

$$\mathbf{P}[t' \le T_0 \le t'', X_{t_1} = x_1, \ldots, X_{t_n} = x_n] = \frac{1}{2}\int_{t'}^{t''} \prod_{k=1}^{n-1} \frac{1+(-1)^{|x_k-x_{k+1}|}e^{-(t_{k+1}-t_k)/t_0}}{2} g(t_0)\mathrm{d}t_0.$$

Under this assumption,

$$\mathbf{P}[X_t = X_{t+\tau}] = \mathbf{P}[X_{t-\tau} = X_t] = \frac{1}{2}\int_{-\infty}^{\infty} g(t_0)(1+e^{-\tau/t_0})\mathrm{d}t_0$$

and

$$\mathbf{P}[X_t = -X_{t+\tau}] = \mathbf{P}[X_{t-\tau} = -X_t] = \frac{1}{2}\int_{-\infty}^{\infty} g(t_0)(1-e^{-\tau/t_0})\mathrm{d}t_0,$$

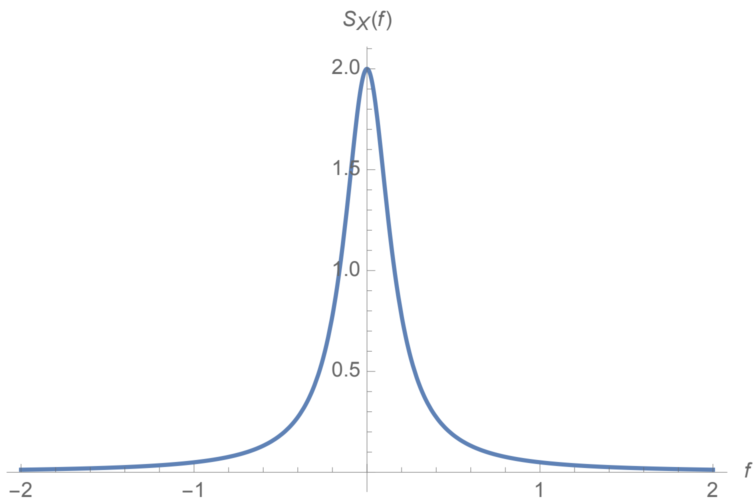

Figure 7: The Lorentzian power spectrum, shown here for $t_0 = 1$.

which gives the autocorrelation function

$$R_X(\tau) = \int_{-\infty}^{\infty} g(t_0) e^{-|\tau|/t} \mathrm{d}t_0$$

and the power spectrum

$$S_X(f) = \int_{-\infty}^{\infty} \frac{2t_0 g(t_0)}{1 + (2\pi t_0 f)^2} \mathrm{d}t_0. \tag{5.27}$$

We will now show that the $1/f$ power spectrum results if we choose the pdf g appropriately.

From semiconductor physics, we know that, in thermal equilibrium at absoulte temperature T, the relaxation time in a material with the energy difference ΔE between the electron and the hole energy levels is given by $ce^{\Delta E/kT}$ for some $c > 0$, where k is the Boltzmann constant. If we assume that the energy gap ΔE is a random variable with a uniform distribution over some range $[\Delta_0, \Delta_1]$, then the pdf g of the relaxation time will be given by

$$g(t_0) = \begin{cases} \dfrac{kT}{(\Delta_1 - \Delta_0)t_0}, & \text{for } ce^{\Delta_0/kT} \le t_0 \le ce^{\Delta_1/kT} \\ 0, & \text{otherwise} \end{cases}. \tag{5.28}$$

Substituting (5.28) into (5.27), we get

$$\begin{aligned} S_X(f) &= \frac{kT}{(\Delta_1 - \Delta_0)} \int_{ce^{\Delta_0/kT}}^{ce^{\Delta_1/kT}} \frac{2}{1 + (2\pi t f)^2} \mathrm{d}t \\ &= \frac{kT}{2\pi f(\Delta_1 - \Delta_0)} \arctan(2\pi f t) \Big|_{ce^{\Delta_0/kT}}^{ce^{\Delta_1/kT}}, \end{aligned}$$

which behaves like $1/f$ for large values of $\Delta_1 - \Delta_0$.