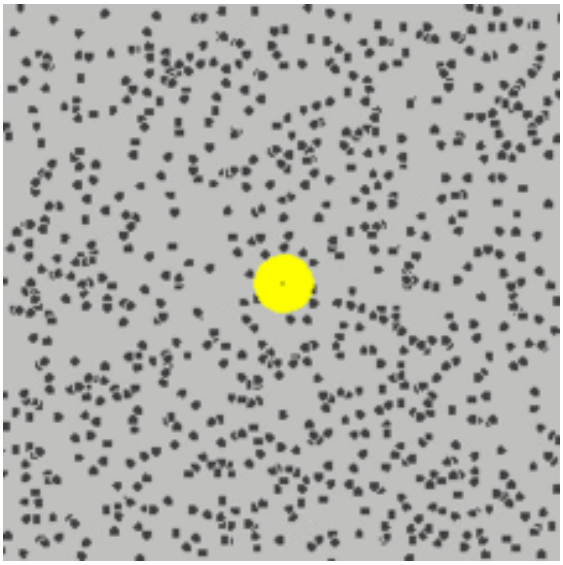


Random Processes

These motions were such as to satisfy me, after frequently repeated observation, that they arose neither from currents in the fluid, nor from its gradual evaporation, but belonged to the particle itself.

ROBERT BROWN (1828)

165

165

6.1 Overview

In this chapter we analyze, among others, the following issues:

- What is the time evolution of the distribution function for an ensemble of systems that begins out of statistical equilibrium and is brought to equilibrium through contact with a heat bath?
- How can one characterize the noise introduced into experiments or observations by noisy devices, such as resistors and amplifiers?
- What is the influence of such noise on one's ability to detect weak signals?
- What filtering strategies will improve one's ability to extract weak signals from strong noise?
- Frictional damping of a dynamical system generally arises from coupling to many other degrees of freedom (a bath) that can sap the system's energy. What is the connection between the fluctuating (noise) forces that the bath exerts on the system and its damping influence?

166

166

6.2.1 Random Variables and Random Processes

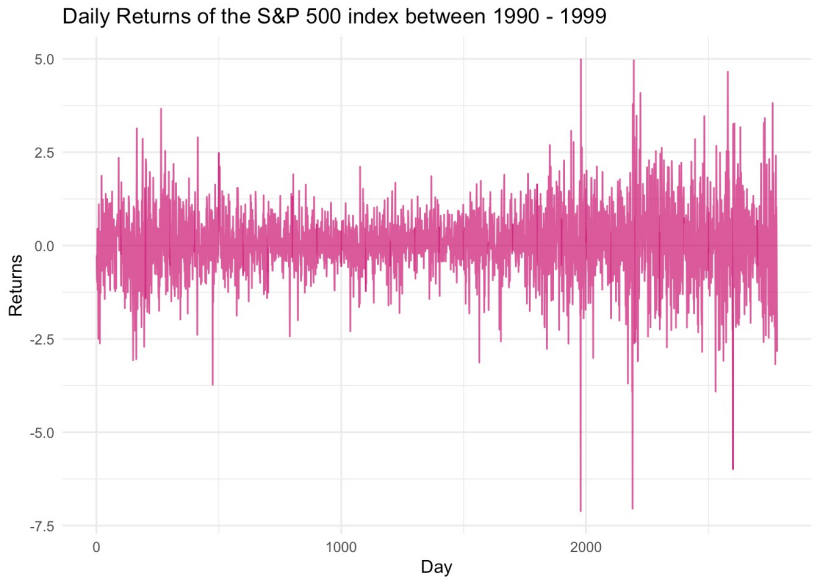
RANDOM VARIABLE

A (1-dimensional) *random variable* is a (scalar) function $y(t)$, where t is usually time, for which the future evolution is not determined uniquely by any set of initial data—or at least by any set that is knowable to you and me. In other words, *random variable* is just a fancy phrase that means "unpredictable function." Throughout this chapter, we insist for simplicity that our random variables y take on a continuum of *real* values ranging over some interval, often but not always $-\infty$ to $+\infty$. The generalizations to y with complex or discrete (e.g., integer) values, and to independent variables other than time, are straightforward.

Examples of random variables are: (i) the total energy $E(t)$ in a cell of gas that is in contact with a heat bath; (ii) the temperature $T(t)$ at the corner of Main Street and Center Street in Logan, Utah; (iii) the price per share of Google stock $P(t)$; (iv) the mass-flow rate $\dot{M}(t)$ from the Amazon River into the Atlantic Ocean. One can also deal with random variables that are vector or tensor functions of time; in Track-Two portions of this chapter we do so.

167

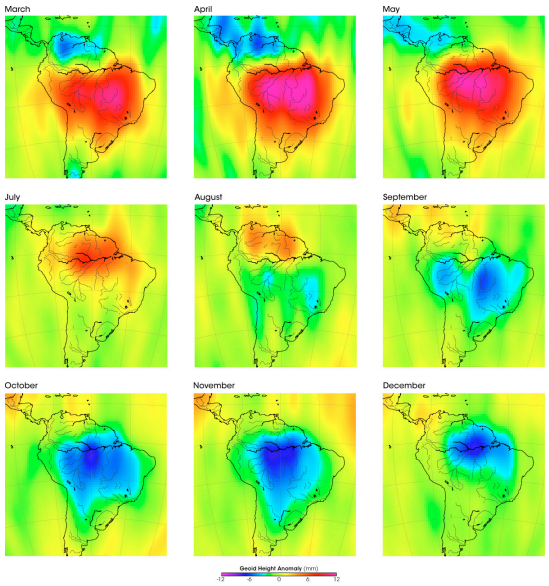
167



168

168

Mass flow rate from the Amazon's river



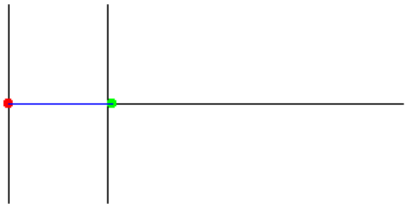
169

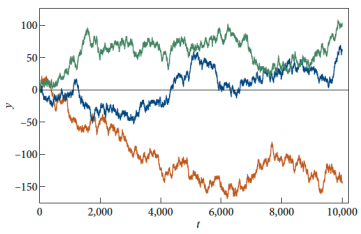
169

RANDOM PROCESS

A (1-dimensional) *random process* (also called "stochastic process") is an ensemble $\mathcal{E}$ of real random variables $y(t)$ that, in a physics context, all represent the same kind of physical entity. For example, each $y(t)$ could be the longitude of a particular oxygen molecule undergoing a random walk in Earth's atmosphere. The individual random variables $y(t)$ in the ensemble $\mathcal{E}$ are often called *realizations* of the random process.

As an example, Fig. 6.1 shows three realizations $y(t)$ of a random process that represents the random walk of a particle in one dimension. For details, see Ex. 6.4, which shows how to generate realizations like these on a computer.

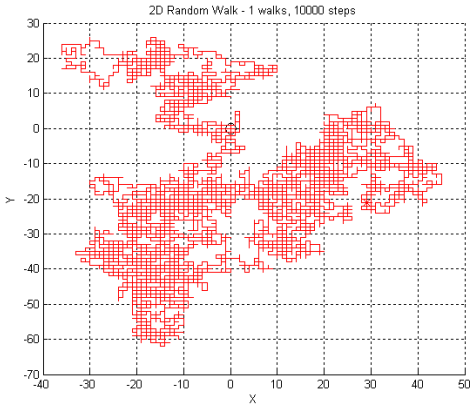


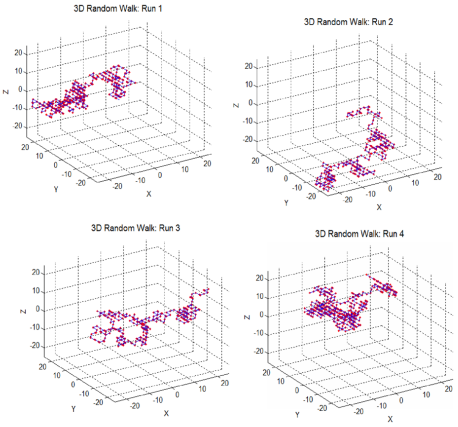


170

170

Random walks in 2d and 3d





171

171

PROBABILITY DISTRIBUTIONS FOR A RANDOM PROCESS

Since the precise time evolution of a random variable $y(t)$ is not predictable, if one wishes to make predictions, one can do so only probabilistically. The foundation for probabilistic predictions is a set of probability functions for the random process (i.e., for the ensemble $\mathcal{E}$ of its realizations).

More specifically, the most general (1-dimensional) random process is fully characterized by the set of probability distributions p_1, p_2, p_3, . . . defined as

$$p_n(y_n, t_n; \ldots; y_2, t_2; y_1, t_1)dy_n \ldots dy_2 dy_1. \qquad (6.1)$$

Equation (6.1) tells us the probability that a realization $y(t)$, drawn at random from the process (the ensemble $\mathcal{E}$), (i) will take on a value between y_1 and $y_1 + dy_1$ at time t_1, (ii) also will take on a value between y_2 and $y_2 + dy_2$ at a later time t_2, . . . , and (iii) also will take on a value between y_n and $y_n + dy_n$ at a later time t_n. (Note that the subscript n on p_n tells us how many independent values of y appear in p_n, and that earlier times are placed to the right—a practice common for physicists, particularly when dealing with propagators.) If we knew the values of all the process's probability distributions (an infinite number of p_ns!), then we would have full information about its statistical properties. Not surprisingly, it will turn out that, if the process in some sense is in statistical equilibrium, then we can compute all its probability distributions from a very small amount of information. But that comes later; first we must develop more formalism.

172

172

ENSEMBLE AVERAGES

From the probability distributions, we can compute ensemble averages (denoted by brackets). For example, the quantities

$$\langle y(t_1)\rangle \equiv \int y_1 p_1(y_1, t_1) dy_1 \quad \text{and} \quad \sigma_y^2(t_1) \equiv \left\langle [y(t_1) - \langle y(t_1)\rangle]^2 \right\rangle \tag{6.2a}$$

are the ensemble-averaged value of y and the variance of y at time t_1. Similarly,

$$\langle y(t_2) y(t_1)\rangle \equiv \int y_2 y_1 p_2(y_2, t_2; y_1, t_1) dy_2 dy_1 \tag{6.2b}$$

is the average value of the product $y(t_2)y(t_1)$.

CONDITIONAL PROBABILITIES

Besides the (absolute) probability distributions p_n, we also find useful an infinite series of *conditional* probability distributions $P_2, P_3, \ldots$, defined as

$$P_n(y_n, t_n | y_{n-1}, t_{n-1}; \ldots; y_1, t_1) dy_n. \tag{6.3}$$

This distribution is the probability that, *if* $y(t)$ took on the values $y_1, y_2, \ldots, y_{n-1}$ at times $t_1, t_2, \ldots, t_{n-1}$, then it will take on a value between y_n and $y_n + dy_n$ at a later time t_n.

173

173

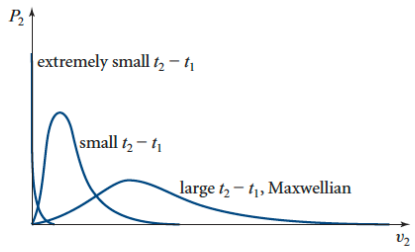


FIGURE 6.2 The probability $P_2(v_2, t_2|0, t_1)$ that a molecule with vanishing speed at time t_1 will have speed v_2 (in a unit interval dv_2) at time t_2. Although the molecular speed is a stationary random process, this probability evolves in time.

STATIONARY RANDOM PROCESSES

A random process is said to be *stationary* if and only if its probability distributions p_n depend just on time differences and not on absolute time:

$$p_n(y_n, t_n + \tau; \ldots; y_2, t_2 + \tau; y_1, t_1 + \tau) = p_n(y_n, t_n; \ldots; y_2, t_2; y_1, t_1). \tag{6.5}$$

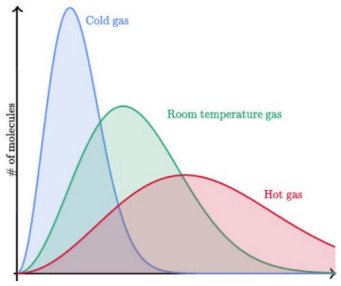
If this property holds for the absolute probabilities p_n, then Eq. (6.4) guarantees it also will hold for the conditional probabilities P_n.

Nonstationary random processes arise when one is studying a system whose evolution is influenced by some sort of clock that registers absolute time, not just time differences. For example, the speeds $v(t)$ of all oxygen molecules in downtown St. Anthony, Idaho, make up random processes regulated in part by the atmospheric temperature and therefore by the rotation of Earth and its orbital motion around the Sun. The influence of these clocks makes $v(t)$ a nonstationary random process. Stationary random processes, by contrast, arise in the absence of any regulating clocks. An example is the speeds $v(t)$ of all oxygen molecules in a room kept at constant temperature.

174

174

Stationary and non-stationary molecular velocity distributions



175

175

Henceforth, throughout this chapter, we restrict attention to random processes that are stationary (at least on the timescales of interest to us); and, accordingly, we use

$$p_1(y) \equiv p_1(y, t_1) \tag{6.6a}$$

for the probability, since it does not depend on the time t_1. We also denote by

$$P_2(y_2, t|y_1) \equiv P_2(y_2, t|y_1, 0) \tag{6.6b}$$

the probability that, if a (realization of a) random process begins with the value y_1, then after the lapse of time t it has the value y_2.

176

176

Ergodic hypothesis

6.2.3 Ergodic Hypothesis

A (stationary) random process (ensemble $\mathcal{E}$ of random variables) is said to satisfy the *ergodic hypothesis* (or, for brevity, it will be called *ergodic*) if and only if it has the following property.

Let $y(t)$ be a random variable in the ensemble $\mathcal{E}$ (i.e., let $y(t)$ be any realization of the process). Construct from $y(t)$ a new ensemble $\mathcal{E}'$ whose members are

$$Y^K(t) \equiv y(t + KT), \tag{6.7}$$

where K runs over all integers, negative and positive, and where T is some very large time interval. Then $\mathcal{E}'$ has the same probability distributions p_n as $\mathcal{E}$; that is, $p_n(Y_n, t_n; \ldots; Y_1, t_1)$ has the same functional form as $p_n(y_n, t_n; \ldots; y_1, t_1)$ for all times such that $|t_i - t_j| < T$.

This is essentially the same ergodic hypothesis as we met in Sec. 4.6.

177

177

Ergodic processes

As in Sec. 4.6, because of the ergodic hypothesis, time averages defined using any realization $y(t)$ of a random process are equal to ensemble averages:

$$\bar{F} \equiv \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} F\big(y(t)\big)dt = \langle F(y)\rangle \equiv \int F(y)p_1(y)dy, \tag{6.8}$$

for any function $F = F(y)$. In this sense, each realization of the random process is representative, when viewed over sufficiently long times, of the statistical properties of the process's entire ensemble—and conversely. Correspondingly, we can blur the distinction between the random process and specific realizations of it—and we often do so.

178

178

6.3.1 Markov Processes; Random Walk

A random process $y(t)$ is said to be *Markov* (also sometimes called "Markovian") if and only if all of its future probabilities are determined by its most recently known value:

$$P_n(y_n, t_n | y_{n-1}, t_{n-1}; \ldots; y_1, t_1) = P_2(y_n, t_n | y_{n-1}, t_{n-1}) \quad \text{for all } t_n \geq \ldots \geq t_2 \geq t_1. \tag{6.9}$$

This relation guarantees that any Markov process (which, of course, we require to be stationary without saying so) is completely characterized by the probabilities

$$p_1(y) \text{ and } P_2(y_2, t | y_1) \equiv \frac{p_2(y_2, t; y_1, 0)}{p_1(y_1)} . \tag{6.10}$$

179

179

An example of a Markov process is the x component of velocity $v_x(t)$ of a dust particle in an arbitrarily large room,[1] filled with constant-temperature air. Why? Because the molecule's equation of motion is[2] $m dv_x/dt = F'_x(t)$, and the force $F'_x(t)$ is due to random buffeting by other molecules that are uncorrelated (the kick now is unrelated to earlier kicks); thus, there is no way for the value of v_x in the future to be influenced by any earlier values of v_x except the most recent one.

By contrast, the position $x(t)$ of the particle is not Markov, because the probabilities of future values of x depend not just on the initial value of x, but also on the initial velocity v_x—or, equivalently, the probabilities depend on the values of x at two initial, closely spaced times. The pair $\{x(t), v_x(t)\}$ is a 2-dimensional Markov process (see Ex. 6.23).

180

180

State Space	Type of Parameter: Discrete	Type of Parameter: Continuous
Discrete	(Discrete-parameter) Markov chain	Continuous-parameter Markov chain
Continuous	Discrete-parameter Markov process	Continuous-parameter Markov process

181

181

THE SMOLUCHOWSKI EQUATION

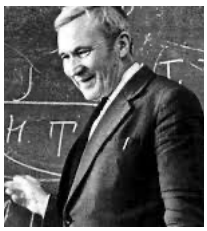
Choose three (arbitrary) times t_1, t_2, and t_3 that are ordered, so $t_1 < t_2 < t_3$. Consider a (realization of an) arbitrary random process that begins with a known value y_1 at t_1, and ask for the probability $P_2(y_3, t_3|y_1)$ (per unit y_3) that it will be at y_3 at time t_3. Since the realization must go through some value y_2 at the intermediate time t_2 (though we don't care what that value is), it must be possible to write the probability to reach y_3 as

$$P_2(y_3, t_3|y_1, t_1) = \int P_3(y_3, t_3|y_2, t_2; y_1, t_1) P_2(y_2, t_2|y_1, t_1) dy_2,$$

where the integration is over all allowed values of y_2. This is not a terribly interesting relation. Much more interesting is its specialization to the case of a Markov process. In that case $P_3(y_3, t_3|y_2, t_2; y_1, t_1)$ can be replaced by $P_2(y_3, t_3|y_2, t_2) = P_2(y_3, t_3 - t_2|y_2, 0) \equiv P_2(y_3, t_3 - t_2|y_2)$, and the result is an integral equation involving only P_2. Because of stationarity, it is adequate to write that equation for the case $t_1 = 0$:

$$\boxed{P_2(y_3, t_3|y_1) = \int P_2(y_3, t_3 - t_2|y_2) P_2(y_2, t_2|y_1) dy_2.} \qquad (6.11)$$

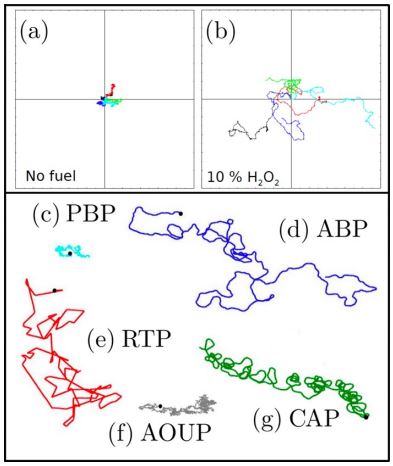
This is the *Smoluchowski equation* (also called *Chapman-Kolmogorov equation*). It is valid for any Markov random process and for times $0 < t_2 < t_3$. We shall discover its power in our derivation of the Fokker-Planck equation in Sec. 6.9.1.



182

182

Passive and active brownian motion



183

183

Models

DRY ACTIVE PARTICLE MODELS FOR A SINGLE PARTICLE			
Model	**Equations of motion**	**Parameters**	**Natural units**
ABP	$\dot{\vec{r}}(t) = \vec{p}(t) + \mathrm{Pe}^{-1}\vec{\xi}(t)$ $\dot{\phi}(t) = \sqrt{2}\eta(t)$	$\mathrm{Pe} = \frac{v_0}{\sqrt{2DD_R}}$	Time scale: $\tau_P = D_R^{-1}$ Length scale: $l = l_P = v_0 D_R^{-1}$
RTP	$\dot{\vec{r}}(t) = \vec{p}(t)$ $\dot{\phi}(t) = \sum_n \Delta\phi_n \delta(t - \vec{T}_n)$	None[a]	Time scale: $\tau_P = \lambda_t^{-1}$ Length scale: $l = l_P = v_0 \lambda_t^{-1}$
AOUP	$\dot{\vec{r}}(t) = \vec{v}_0(t)$ $\dot{\vec{v}}_0(t) = -\vec{v}_0(t) + \sqrt{2}\vec{\xi}(t)$	None	Time scale: τ_P Length scale: $l = \sqrt{D\tau_P}$
CAP	$\dot{\vec{r}}(t) = \vec{p}(t) + \mathrm{Pe}^{-1}\vec{\xi}(t)$ $\dot{\phi}(t) = \tilde{\omega} + \sqrt{2}\eta(t)$	$\mathrm{Pe} = \frac{v_0}{\sqrt{2DD_R}}$ $\tilde{\omega} = \omega\tau_P$	Time scale: $\tau_P = D_R^{-1}$ Length scale: $l = l_P = v_0 D_R^{-1}$

184

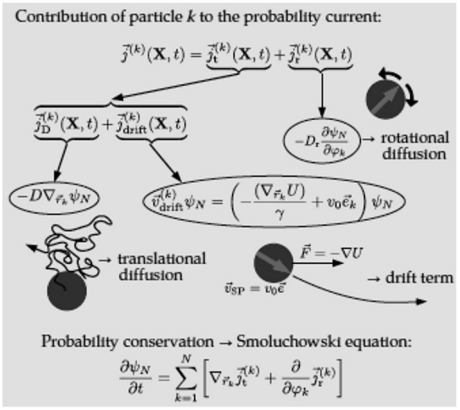
184

Smoluchowski equation for ABPs (new project)

$$\frac{\partial \psi_N}{\partial t} = \sum_{k=1}^{N} \nabla_{\vec{r}_k} \cdot \left[\frac{(\nabla_{\vec{r}_k} U)}{\gamma} - v_0 \vec{p}_k + D \nabla_{\vec{r}_k} \right] \psi_N + D_R \sum_{k=1}^{N} \frac{\partial^2 \psi_N}{\partial \varphi_k^2},$$

An Introduction to Modeling Approaches of Active Matter

L. Hecht, J. C. Ureña, and B. Liebchen*
Institut für Physik kondensierter Materie, Technische Universität Darmstadt, Hochschulstr. 8, 64289 Darmstadt, Germany



185

185

GAUSSIAN PROCESSES

A random process is said to be Gaussian if and only if all of its (absolute) probability distributions are Gaussian (i.e., have the following form):

$$p_n(y_n, t_n; \ldots; y_2, t_2; y_1, t_1) = A \exp\left[-\sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} (y_j - \bar{y})(y_k - \bar{y}) \right], \quad (6.14a)$$

where (i) A and α_{jk} depend only on the time differences $t_2 - t_1, t_3 - t_1, \ldots, t_n - t_1$; (ii) A is a positive normalization constant; (iii) $[\alpha_{jk}]$ is a *positive-definite*, symmetric matrix (otherwise p_n would not be normalizable); and (iv) $\bar{y}$ is a constant, which one readily can show is equal to the ensemble average of y,

$$\bar{y} \equiv \langle y \rangle = \int y p_1(y)\, dy. \quad (6.14b)$$

Since the conditional probabilities are all computable as ratios of absolute probabilities [Eq. (6.4)], the conditional probabilities of a Gaussian process will be Gaussian.

186

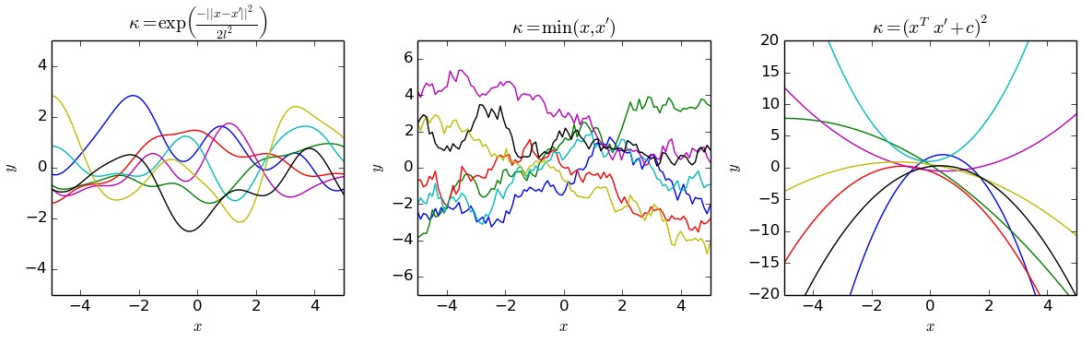
186

Gaussian random processes are very common in physics. For example, the total number of particles $N(t)$ in a gas cell that is in statistical equilibrium with a heat bath is a Gaussian random process (Ex. 5.11d); and the primordial fluctuations that gave rise to structure in our universe appear to have been Gaussian (Sec. 28.5.3). In fact, as we saw in Sec. 5.6, macroscopic variables that characterize huge systems in statistical equilibrium always have Gaussian probability distributions. The underlying reason is that, *when a random process is driven by a large number of statistically independent, random influences, its probability distributions become Gaussian*. This general fact is a consequence of the *central limit theorem* of probability. We state and prove a simple variant of this theorem.

187

187

Gaussian processes with different kernels



188

188

6.3.3 Doob's Theorem for Gaussian-Markov Processes, and Brownian Motion

A large fraction of the random processes that one meets in physics are Gaussian, and many are Markov. Therefore, the following remarkable theorem is very important. *Any 1-dimensional random process $y(t)$ that is both Gaussian and Markov has the following form for its conditional probability distribution P_2:*

$$P_2(y_2, t|y_1) = \frac{1}{[2\pi {\sigma_{y_t}}^2]^{\frac{1}{2}}} \exp\left[-\frac{(y_2 - \bar{y}_t)^2}{2{\sigma_{y_t}}^2}\right], \tag{6.18a}$$

where the mean $\bar{y}_t$ and variance $\sigma^2_{y_t}$ at time t are given by

$$\bar{y}_t = \bar{y} + e^{-t/\tau_r}(y_1 - \bar{y}), \quad \sigma^2_{y_t} = (1 - e^{-2t/\tau_r}){\sigma_y}^2. \tag{6.18b}$$

Here $\bar{y}$ and ${\sigma_y}^2$ are respectively the process's equilibrium mean and variance (the values at $t \to \infty$), and τ_r is its *relaxation time*. This result is *Doob's theorem*.[5] We shall prove it in Ex. 6.5, after we have developed some necessary tools.

189

189

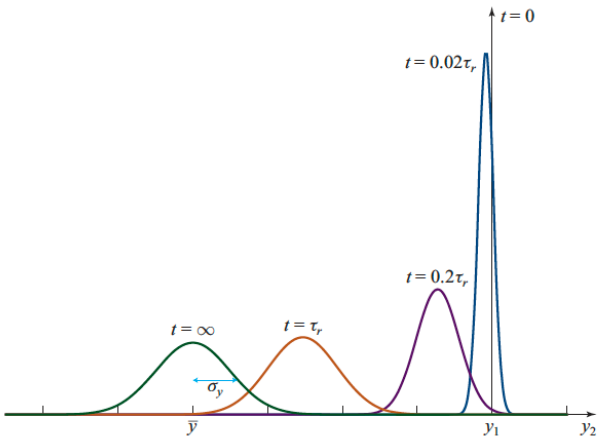


FIGURE 6.4 Evolution of the conditional probability $P_2(y_2, t|y_1)$ for a Gaussian-Markov random process [Eq. (6.18a)], as predicted by Doob's theorem. The correlation function and spectral density for this process are shown later in the chapter in Fig. 6.8.

190

190

6.4 Correlation Functions and Spectral Densities

6.4.1 Correlation Functions;

Let $y(t)$ be a (realization of a) random process with time average $\bar{y}$. Then the correlation function of $y(t)$ is defined by

$$C_y(\tau) \equiv \overline{[y(t)-\bar{y}][y(t+\tau)-\bar{y}]} \equiv \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{+T/2} [y(t)-\bar{y}][y(t+\tau)-\bar{y}]dt. \tag{6.19}$$

This quantity, as its name suggests, is a measure of the extent to which the values of y at times t and $t+\tau$ tend to be correlated. The quantity τ is sometimes called the *delay time*, and by convention it is taken to be positive. [One can easily see that, if one also defines $C_y(\tau)$ for negative delay times τ by Eq. (6.19), then $C_y(-\tau) = C_y(\tau)$. Thus nothing is lost by restricting attention to positive delay times.]

As an example, for a Gaussian-Markov process with P_2 given by Doob's formula (6.18a) (Fig. 6.4), we can compute $C(\tau)$ by replacing the time average in Eq. (6.19) with an ensemble average: $C_y(\tau) = \int y_2\, y_1\, p_2(y_2, \tau; y_1)\, dy_1\, dy_2$. If we use $p_2(y_2, \tau; y_1) = P_2(y_2, \tau; y_1)\, p_1(y_1)$ [Eq. (6.10)], insert P_2 and p_1 from Eqs. (6.18), and perform the integrals, we obtain

$$C_y(\tau) = \sigma_y^2 e^{-\tau/\tau_r}. \tag{6.20}$$

This correlation function has two properties that are quite general:

1. The following is true for all (ergodic and stationary) random processes:

$$C_y(0) = \sigma_y^2, \tag{6.21a}$$

as one can see by replacing time averages with ensemble averages in definition (6.19); in particular, $C_y(0) \equiv \overline{(y-\bar{y})^2} = \langle (y-\bar{y})^2 \rangle$, which by definition is the variance σ_y^2 of y.

2. In addition, we have that

$$C_y(\tau) \text{ asymptotes to zero for } \tau \gg \tau_r, \tag{6.21b}$$

where τ_r is the process's *relaxation time* or *correlation time* (see Fig. 6.5). This is true for all ergodic, stationary random processes, since our definition of ergodicity in Sec. 6.2.3 relies on each realization $y(t)$ losing its memory of earlier values after some sufficiently long time T. Otherwise, it would not be possible to construct the ensemble $\mathcal{E}'$ of random variables $Y^K(t)$ [Eq. (6.7)] and have them behave like independent random variables.

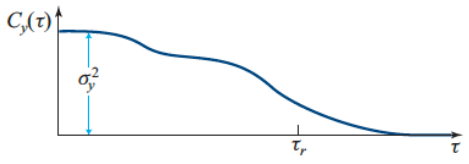


FIGURE 6.5 Properties (6.21) of correlation functions.

6.4.2 Spectral Densities

There are several different normalization conventions for Fourier transforms. In this chapter, we adopt a normalization that is commonly (though not always) used in the theory of random processes and that differs from the one common in quantum theory. Specifically, instead of using the angular frequency ω, we use the ordinary frequency $f \equiv \omega/(2\pi)$. We define the Fourier transform of a function $y(t)$ and its inverse by

$$\boxed{\tilde{y}(f) \equiv \int_{-\infty}^{+\infty} y(t) e^{i2\pi f t} dt, \qquad y(t) \equiv \int_{-\infty}^{+\infty} \tilde{y}(f) e^{-i2\pi f t} df.} \tag{6.23}$$

Notice that with this set of conventions, there are no factors of $1/(2\pi)$ or $1/\sqrt{2\pi}$ multiplying the integrals. Those factors have been absorbed into the df of Eq. (6.23), since $df = d\omega/(2\pi)$.

The integrals in Eq. (6.23) are not well defined as written because a random process $y(t)$ is generally presumed to go on forever so its Fourier transform $\tilde{y}(f)$ is divergent. One gets around this problem by crude trickery. From $y(t)$ construct, by truncation, the function

$$y_T(t) \equiv \begin{cases} y(t) & \text{if } -T/2 < t < +T/2, \\ 0 & \text{otherwise.} \end{cases} \tag{6.24a}$$

Then the Fourier transform $\tilde{y}_T(f)$ is finite, and by Parseval's theorem (e.g., Arfken, Weber, and Harris, 2013) it satisfies

$$\int_{-T/2}^{+T/2} [y(t)]^2 dt = \int_{-\infty}^{+\infty} [y_T(t)]^2 dt = \int_{-\infty}^{+\infty} |\tilde{y}_T(f)|^2 df = 2\int_0^{\infty} |\tilde{y}_T(f)|^2 \, df. \tag{6.24b}$$

In the last equality we have used the fact that because $y_T(t)$ is real, $\tilde{y}_T^*(f) = \tilde{y}_T(-f)$, where * denotes complex conjugation. Consequently, the integral from $-\infty$ to 0 of $|\tilde{y}_T(f)|^2$ is the same as the integral from 0 to $+\infty$. Now, the quantities on the two sides of (6.24b) diverge in the limit as $T \to \infty$, and it is obvious from the left-hand side that they diverge linearly as T. Correspondingly, the limit

$$\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} [y(t)]^2 dt = \lim_{T\to\infty} \frac{2}{T} \int_0^{\infty} |\tilde{y}_T(f)|^2 df \tag{6.24c}$$

is convergent.

These considerations motivate the following definition of the *spectral density* (also sometimes called the *power spectrum*) $S_y(f)$ of the random process $y(t)$:

$$\boxed{S_y(f) \equiv \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{+T/2} [y(t) - \bar{y}] e^{i2\pi f t} dt \right|^2.} \tag{6.25}$$

Notice that the quantity inside the absolute value sign is just $\tilde{y}_T(f)$, but with the mean of y removed before computation of the Fourier transform. (The mean is removed to avoid an uninteresting delta function in $S_y(f)$ at zero frequency.) Correspondingly, by virtue of our motivating result (6.24c), the spectral density satisfies $\int_0^\infty S_y(f)df = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{+T/2}[y(t)-\bar{y}]^2 dt = \overline{(y-\bar{y})^2} = \sigma_y{}^2$, or

$$\boxed{\int_0^\infty S_y(f)df = C_y(0) = \sigma_y^2.} \tag{6.26}$$

Thus the integral of the spectral density of y over all positive frequencies is equal to the variance of y.

By convention, our spectral density is defined only for nonnegative frequencies f. This is because, were we to define it also for negative frequencies, the fact that $y(t)$ is real would imply that $S_y(f) = S_y(-f)$, so the negative frequencies contain no new information. Our insistence that f be positive goes hand in hand with the factor 2 in the $2/T$ of definition (6.25): that factor 2 folds the negative-frequency part onto the positive-frequency part. This choice of convention is called the *single-sided spectral density*. Sometimes one encounters a *double-sided spectral density*,

$$S_y^{\text{double-sided}}(f) = \frac{1}{2}S_y(|f|), \tag{6.27}$$

in which f is regarded as both positive and negative, and frequency integrals generally run from $-\infty$ to $+\infty$ instead of 0 to ∞ (see, e.g., Ex. 6.7).

195

195

Meaning of the spectral density

We can infer the physical meaning of the spectral density from previous experience with light spectra. Specifically, consider the scalar electric field[6] $E(t)$ of a plane-polarized light wave entering a telescope from a distant star, galaxy, or nebula. (We must multiply this $E(t)$ by the polarization vector to get the vectorial electric field.) This $E(t)$ is a superposition of emission from an enormous number of atoms, molecules, and high-energy particles in the source, so it is a Gaussian random process. It is not hard to convince oneself that $E(t)$'s spectral density $S_E(f)$ is proportional to the light power per unit frequency $d\mathcal{E}/dtdf$ (the light's power spectrum) entering the telescope. When we send the light through a diffraction grating, we get this power spectrum spread out as a function of frequency f in the form of spectral lines superposed on a continuum, as in Fig. 6.6. The amount of light power in this spectrum, in some narrow bandwidth Δf centered on some frequency f, is $(d\mathcal{E}/dtdf)\Delta f \propto S_E(f)\Delta f$ (assuming S_E is nearly constant over that band).

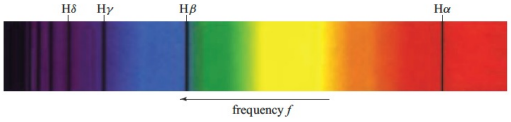
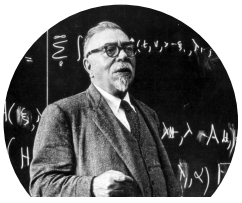


FIGURE 6.6 A spectrum obtained by sending light through a diffraction grating. The intensity of the image is proportional to $d\mathcal{E}/dtdf$, which, in turn, is proportional to the spectral density $S_E(f)$ of the electric field $E(t)$ of the light that entered the diffraction grating.

196

196



6.4.4 The Wiener-Khintchine Theorem;

The Wiener-Khintchine theorem says that, *for any random process $y(t)$, the correlation function $C_y(\tau)$ and the spectral density $S_y(f)$ are the cosine transforms of each other and thus contain precisely the same information:*

$$C_y(\tau) = \int_0^\infty S_y(f)\cos(2\pi f\tau)df, \qquad S_y(f) = 4\int_0^\infty C_y(\tau)\cos(2\pi f\tau)d\tau. \tag{6.29}$$

The factor 4 results from our folding negative frequencies into positive in our definition of the spectral density.

197

Proof of Wiener-Khintchine Theorem. This theorem is readily proved as a consequence of Parseval's theorem: Assume, from the outset, that the mean has been subtracted from $y(t)$, so $\bar{y} = 0$. (This is not really a restriction on the proof, since C_y and S_y are insensitive to the mean of y.) Denote by $y_T(t)$ the truncated y of Eq. (6.24a) and by $\tilde{y}_T(f)$ its Fourier transform. Then the generalization of Parseval's theorem[7]

$$\int_{-\infty}^{+\infty}(gh^* + hg^*)dt = \int_{-\infty}^{+\infty}(\tilde{g}\tilde{h}^* + \tilde{h}\tilde{g}^*)df \tag{6.30a}$$

[with $g = y_T(t)$ and $h = y_T(t+\tau)$ both real and with $\tilde{g} = \tilde{y}_T(f), \tilde{h} = \tilde{y}_T(f)e^{-i2\pi f\tau}$], states

$$\int_{-\infty}^{+\infty} y_T(t)y_T(t+\tau)dt = \int_{-\infty}^{+\infty}\tilde{y}_T^*(f)\tilde{y}_T(f)e^{-i2\pi f\tau}\,df. \tag{6.30b}$$

By dividing by T, taking the limit as $T \to \infty$, and using Eqs. (6.19) and (6.25), we obtain the first equality of Eqs. (6.29). The second follows from the first by Fourier inversion. ■

198

As an application of the Wiener-Khintchine theorem, we can deduce the spectral density $S_y(f)$ for any Gaussian-Markov process by performing the cosine transform of its correlation function $C_y(\tau) = \sigma_y{}^2 e^{-\tau/\tau_r}$ [Eq. (6.20)]. The result is

$$S_y(f) = \frac{(4/\tau_r)\sigma_y{}^2}{(2\pi f)^2 + (1/\tau_r)^2}; \tag{6.32}$$

see Fig. 6.8.

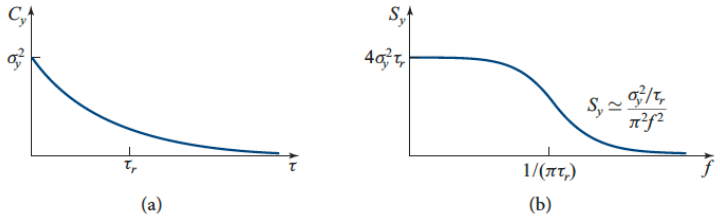


FIGURE 6.8 (a) The correlation function (6.20) and (b) the spectral density (6.32) for a Gaussian-Markov process. The conditional probability $P_2(y_2, \tau | y_1)$ for this process is shown in Fig. 6.4.

199

199

6.6 Noise and Its Types of Spectra

Experimental physicists and engineers encounter random processes in the form of noise that is superposed on signals they are trying to measure. Examples include:

1. In radio communication, static on the radio is noise.
2. When modulated laser light is used for optical communication, random fluctuations in the arrival times of photons always contaminate the signal; the effects of such fluctuations are called "shot noise" and will be studied in Sec. 6.6.1.
3. Even the best of atomic clocks fail to tick with absolutely constant angular frequencies ω. Their frequencies fluctuate ever so slightly relative to an ideal clock, and those fluctuations can be regarded as noise.

200

200

6.6.1 Shot Noise, Flicker Noise, and Random-Walk Noise; Cesium Atomic Clock

Physicists, astronomers, and engineers give names to certain shapes of noise spectra:

$$S_y(f) \text{ independent of } f\text{—white noise spectrum,} \tag{6.44a}$$

$$S_y(f) \propto 1/f\text{—flicker noise spectrum,} \tag{6.44b}$$

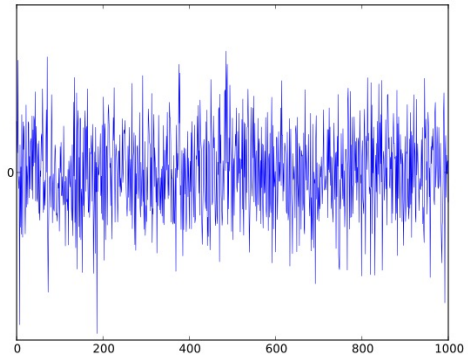
$$S_y(f) \propto 1/f^2\text{—random-walk spectrum.} \tag{6.44c}$$

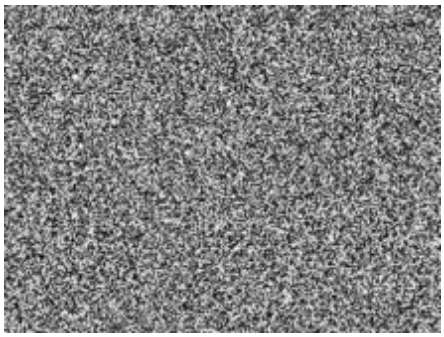
201

201

White noise (f independente)

White noise, S_y independent of f, is called "white" because it has equal amounts of power per unit frequency S_y at all frequencies, just as white light has roughly equal powers at all light frequencies. Put differently, if $y(t)$ has a white-noise spectrum, then its rms fluctuations in fixed bandwidth Δf are independent of frequency f (i.e., $\sqrt{S_y(f)\Delta f}$ is independent of f).





202

202

Flicker noise (1/f)

Flicker noise, $S_y \propto 1/f$, gets its name from the fact that, when one looks at the time evolution $y(t)$ of a random process with a flicker-noise spectrum, one sees fluctuations ("flickering") on all timescales, and the rms amplitude of flickering is independent of the timescale one chooses. Stated more precisely, choose any timescale Δt and then choose a frequency $f \sim 3/\Delta t$, so one can fit roughly three periods of oscillation into the chosen timescale. Then the rms amplitude of the fluctuations observed will be $\sqrt{S_y(f)f/3}$, which is a constant independent of f when the spectrum is that of flicker noise, $S_y \propto 1/f$. In other words, flicker noise has the same amount of power in each octave of frequency. Figure 6.10 is an illustration: both graphs shown there depict random processes with flicker-noise spectra. (The differences between the two graphs will be explained in Sec. 6.6.2.) No matter what time interval one chooses, these processes look roughly periodic with one, two, or three oscillations in that time interval; and the amplitudes of those oscillations are independent of the chosen time interval. Flicker noise occurs widely in the real world, at low frequencies, for instance, in many electronic devices, in some atomic clocks, in geophysics (the flow rates of rivers, ocean currents, etc.), in astrophysics (the light curves of quasars, sunspot numbers, etc.); even in classical music. For an interesting discussion, see Press (1978).

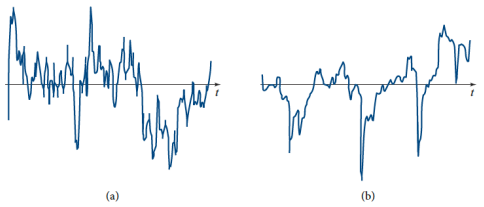
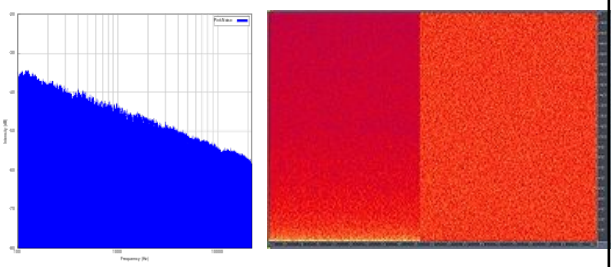


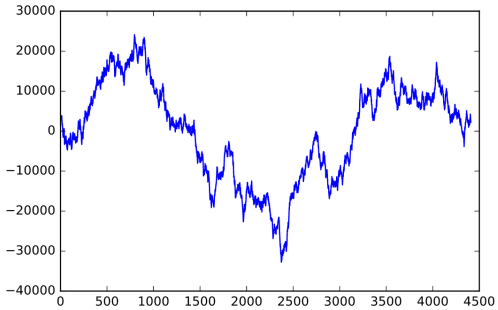
FIGURE 6.10 Examples of two random processes that have flicker noise spectra, $S_y(f) \propto 1/f$. Adapted from Press (1978).

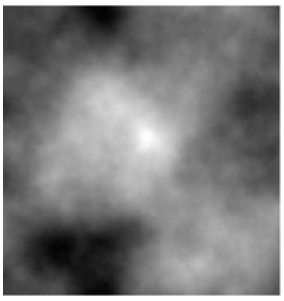


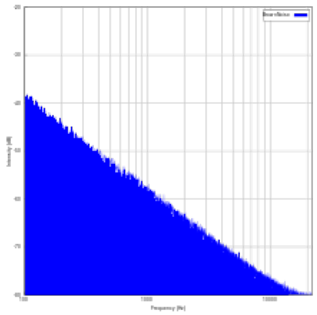
203

203

Brown or random walk noise (1/f²)





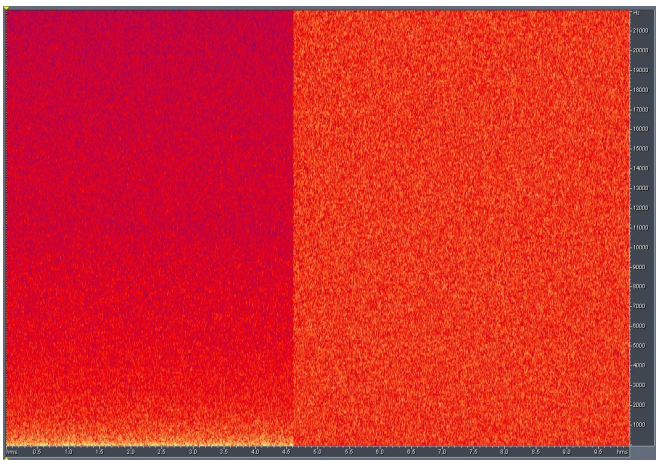


Random-walk noise, $S_y \propto 1/f^2$, arises when a random process $y(t)$ undergoes a random walk. In Sec. 6.7.2, we explore an example: the time evolving position $x(t)$ of a dust particle buffeted by air molecules—the phenomenon of Brownian motion.

204

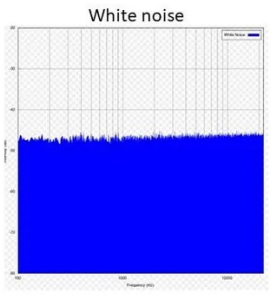
204

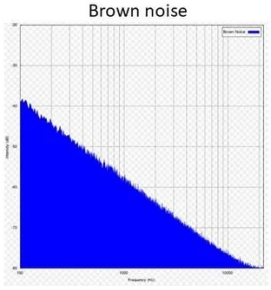
Flicker vs White noise (f vs t)

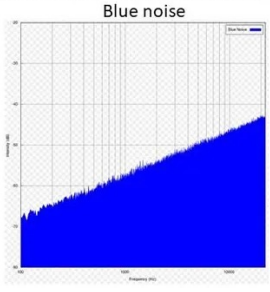


205

205







206

206

Noise in atomic clocks

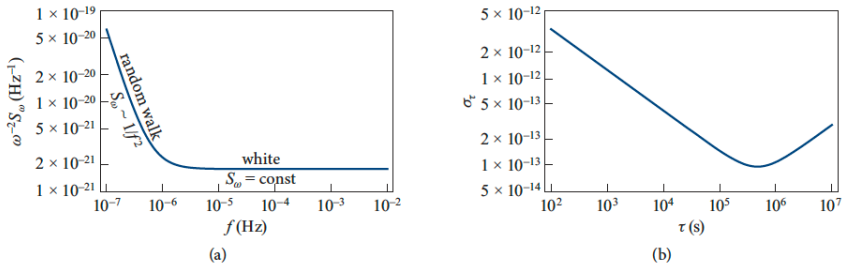
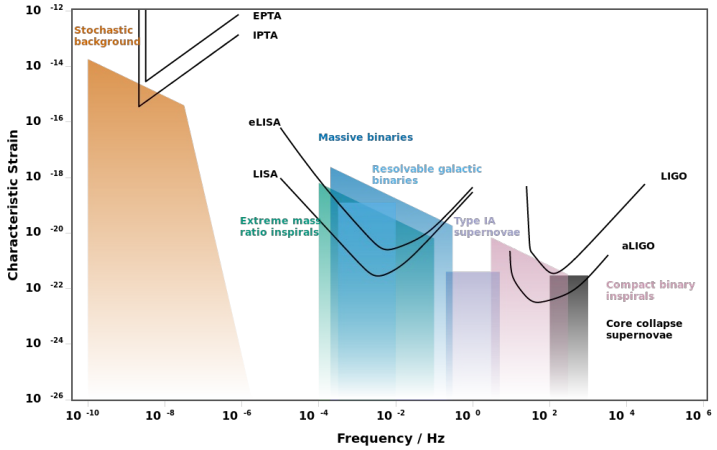


FIGURE 6.11 (a) Spectral density of the fluctuations in angular frequency ω of a typical cesium atomic clock. (b) Square root of the Allan variance for the same clock; see Ex. 6.13. Adapted from Galleani (2012). The best cesium clocks in 2016 (e.g., the U.S. primary time and frequency standard) have amplitude noise, $\sqrt{S_\omega}$ and σ_τ, 1000 times lower than this.

207

207

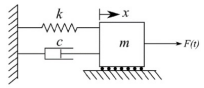
Noise in gravitational-wave detectors as a function of frequency

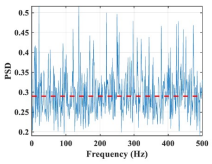


208

208

Fluctuation-dissipation theorem





Friction is generally caused by interaction with the huge number of degrees of freedom of some sort of bath (e.g., the molecules of air against which a moving ball or dust particle pushes). Those degrees of freedom also produce fluctuating forces. In this section, we study the relationship between the friction and the fluctuating forces when the bath is thermalized at some temperature T (so it is a heat bath).

For simplicity, we restrict ourselves to a specific generalized coordinate q of the system that will interact with a bath (e.g., the x coordinate of the ball or dust particle).

We require just one special property for q: its time derivative $\dot{q} = dq/dt$ must appear in the system's lagrangian as a kinetic energy,

$$E_{\text{kinetic}} = \frac{1}{2} m \dot{q}^2, \tag{6.70}$$

and in no other way. Here m is a (generalized) mass associated with q. Then the equation of motion for q will have the simple form of Newton's second law, $m\ddot{q} = F$, where F includes contributions $\mathcal{F}$ from the system itself (e.g., a restoring force in the case of a normal mode), plus a force F_{bath} due to the heat bath (i.e., due to all the degrees of freedom in the bath). This F_{bath} is a random process whose time average is a frictional (damping) force proportional to $\dot{q}$:

$$\bar{F}_{\text{bath}} = -R\dot{q}, \quad F_{\text{bath}} \equiv \bar{F}_{\text{bath}} + F'. \tag{6.71}$$

Here R is the coefficient of friction. The fluctuating part F' of F_{bath} is responsible for driving q toward statistical equilibrium.

209

209

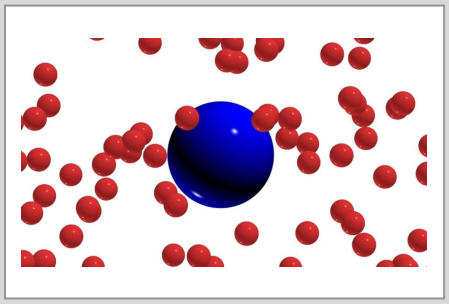
Three specific examples, to which we shall return below, are as follows.

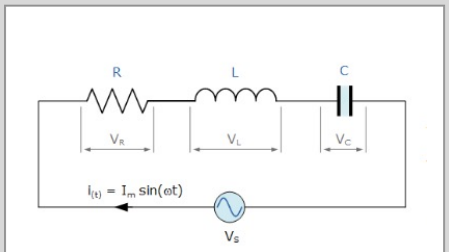
1. The system might be a dust particle with q its x coordinate and m its mass. The heat bath might be air molecules at temperature T, which buffet the dust particle, producing Brownian motion.
2. The system might be an L-C-R circuit (i.e., an electric circuit containing an inductance L, a capacitance C, and a resistance R) with q the total electric charge on the top plate of the capacitor. The bath in this case would be the many mechanical degrees of freedom in the resistor. For such a circuit, the "equation of motion" is

$$L\ddot{q} + C^{-1}q = F_{\text{bath}}(t) = -R\dot{q} + F', \tag{6.72}$$

so the effective mass is the inductance L; the coefficient of friction is the resistance (both denoted R); $-R\dot{q} + F'$ is the total voltage across the resistor; and F' is the fluctuating voltage produced by the resistor's internal degrees of freedom (the bath) and so might better be denoted V'.

3. The system might be the fundamental mode of a 10-kg sapphire crystal with q its generalized coordinate (cf. Sec. 4.2.1). The heat bath might be all the other normal modes of vibration of the crystal, with which the fundamental mode interacts weakly.





210

210

LANGEVIN EQUATION

In general, the equation of motion for the generalized coordinate $q(t)$ under the joint action of (i) the bath's damping force $-R\dot{q}$, (ii) the bath's fluctuating forces F', and (iii) the system's internal force $\mathcal{F}$ will take the form [cf. Eq. (6.71)]

$$m\ddot{q} + R\dot{q} = \mathcal{F} + F'(t). \tag{6.73}$$

The internal force $\mathcal{F}$ is derived from the system's hamiltonian or lagrangian in the absence of the heat bath. For the L-C-R circuit of Eq. (6.72) that force is $\mathcal{F} = -C^{-1}q$; for the dust particle, if the particle were endowed with a charge Q and were in an external electric field with potential $\Phi(t, x, y, z)$, it would be $\mathcal{F} = -Q\partial\Phi/\partial x$; for the normal mode of a crystal, it is $\mathcal{F} = -m\omega^2 q$, where ω is the mode's eigenfrequency.

211

211

Because the equation of motion (6.73) involves a driving force $F'(t)$ that is a random process, one cannot solve it to obtain $q(t)$. Instead, one must solve it in a statistical way to obtain the evolution of q's probability distributions $p_n(q_1, t_1; \ldots; q_n, t_n)$. This and other evolution equations involving random-process driving terms are called by modern mathematicians *stochastic differential equations,* and there is an extensive body of mathematical formalism for solving them. In statistical physics the specific stochastic differential equation (6.73) is known as the *Langevin equation.*

212

212

ELEMENTARY FLUCTUATION-DISSIPATION THEOREM

Because the damping force $-R\dot{q}$ and the fluctuating force F' both arise from interaction with the same heat bath, there is an intimate connection between them. For example, the stronger the coupling to the bath, the stronger will be the coefficient of friction R and the stronger will be F'. The precise relationship between the dissipation embodied in R and the fluctuations embodied in F' is given by the following *fluctuation-dissipation theorem:* At frequencies

$$f \ll 1/\tau_r, \tag{6.74a}$$

where τ_r is the (very short) relaxation time for the fluctuating force F', the fluctuating force has the spectral density

$$\boxed{S_{F'}(f) = 4R\left(\frac{1}{2}hf + \frac{hf}{e^{hf/(k_BT)} - 1}\right) \quad \text{in general,}} \tag{6.74b}$$

$$\boxed{S_{F'}(f) = 4Rk_BT \quad \text{in the classical domain, } k_BT \gg hf.} \tag{6.74c}$$

Here T is the temperature of the bath, and h is Planck's constant.

213

213

Notice that in the classical domain, $k_BT \gg hf$, the spectral density has a white-noise spectrum. In fact, since we are restricting attention to frequencies at which F' has no self-correlations ($f^{-1} \gg \tau_r$), F' is Markov; and since it is produced by interaction with the huge number of degrees of freedom of the bath, F' is also Gaussian. Thus, in the classical domain F' is a Gaussian-Markov, white-noise process.

At frequencies $f \gg k_BT/h$ (quantum domain), in Eq. (6.74b) the term $S_{F'} = 4R\frac{1}{2}hf$ is associated with vacuum fluctuations of the degrees of freedom that make up the heat bath (one-half quantum of fluctuations per mode as for any quantum mechanical simple harmonic oscillator). In addition, the second term, $S_{F'}(f) = 4Rhfe^{-hf/(k_BT)}$, associated with thermal excitations of the bath's degrees of freedom, is exponentially suppressed because at these high frequencies, the bath's modes have exponentially small probabilities of containing any quanta at all. Since in this quantum domain $S_{F'}(f)$ does not have the standard Gaussian-Markov frequency dependence (6.32), in the quantum domain F' is not a Gaussian-Markov process.

214

214

Proof of the Fluctuation-Dissipation Theorem.

In principle, we can alter the system's internal restoring force $\mathcal{F}$ without altering its interactions with the heat bath [i.e., without altering R or $S_{F'}(f)$]. For simplicity, we set $\mathcal{F}$ to zero so q becomes the coordinate of a free mass. The basic idea of our proof is to choose a frequency f_o at which to evaluate the spectral density of F', and then, in an idealized thought experiment, very weakly couple a harmonic oscillator with eigenfrequency f_o to q. Through that coupling, the oscillator is indirectly damped by the resistance R of q and is indirectly driven by R's associated fluctating force F', which arises from a bath with temperature T. After a long time, the oscillator will reach thermal equilibrium with that bath and will then have the standard thermalized mean kinetic energy ($\bar{E} = k_B T$ in the classical regime). We shall compute that mean energy in terms of $S_{F'}(f_o)$ and thereby deduce $S_{F'}(f_o)$.

215

215

The Langevin equation (6.73) and equation of motion for the coupled free mass and harmonic oscillator are

$$m\ddot{q} + R\dot{q} = -\kappa Q + F'(t), \quad M\ddot{Q} + M\omega_o^2 Q = -\kappa q. \tag{6.75a}$$

Here M, Q, and $\omega_o = 2\pi f_o$ are the oscillator's mass, coordinate, and angular eigenfrequency, and κ is the arbitrarily small coupling constant. (The form of the coupling terms $-\kappa Q$ and $-\kappa q$ in the two equations can be deduced from the coupling's interaction hamiltonian $H_I = \kappa q Q$.) Equations (6.75a) can be regarded as a filter to produce from the fluctuating-force input $F'(t)$ a resulting motion of the oscillator, $Q(t) = \int_{-\infty}^{+\infty} K(t-t')F'(t')dt'$. The squared Fourier transform $|\tilde{K}(f)|^2$ of this filter's kernel $K(t-t')$ is readily computed by the standard method [Eq. (6.51) and associated discussion] of inserting a sinusoid $e^{-i\omega t}$ (with $\omega = 2\pi f$) into the filter [i.e., into the differential equations (6.75a)] in place of F', then solving for the sinusoidal output Q, and then setting $|\tilde{K}|^2 = |Q|^2$. The resulting $|\tilde{K}|^2$ is the ratio of the spectral densities of input and output. We carefully manipulate the resulting $|\tilde{K}|^2$ so as to bring it into the following standard resonant form:

$$S_q(f) = |\tilde{K}(f)|^2 S_{F'}(f) = \frac{|B|^2}{(\omega - \omega_o')^2 + (2M\omega_o^2 R|B|^2)^2]} S_{F'}(f). \tag{6.75b}$$

216

216

Here $B = \kappa/[2M\omega_o(m\omega_o^2 + iR\omega_o)]$ is arbitrarily small because κ is arbitrarily small; and $\omega_o'^2 = \omega_o^2 + 4mM\omega_o^4|B|^2$ is the oscillator's squared angular eigenfrequency after coupling to q, and is arbitrarily close to ω_o^2 because $|B|^2$ is arbitrarily small. In these equations we have replaced ω by ω_o everywhere except in the resonance term $(\omega - \omega_o')^2$ because $|\tilde{K}|^2$ is negligibly small everywhere except near resonance, $\omega \cong \omega_o$.

The mean energy of the oscillator, averaged over an arbitrarily long timescale, can be computed in either of two ways.

217

217

1. Because the oscillator is a mode of some boson field and is in statistical equilibrium with a heat bath, its mean occupation number must have the standard Bose-Einstein value $\eta = 1/[e^{\hbar\omega_o/(k_BT)} - 1]$, and since each quantum carries an energy $\hbar\omega_o$, the mean energy is

$$\bar{E} = \frac{\hbar\omega_o}{e^{\hbar\omega_o/(k_BT)} - 1} + \frac{1}{2}\hbar\omega_o. \tag{6.75c}$$

   Here we have included the half-quantum of energy associated with the mode's vacuum fluctuations.

2. Because on average the energy is half potential and half kinetic, and the mean potential energy is $\frac{1}{2}m\omega_o^2\overline{Q^2}$, and because the ergodic hypothesis tells us that time averages are the same as ensemble averages, it must be that

$$\bar{E} = 2\frac{1}{2}M\omega_o^2\omega^2\langle Q^2\rangle = M\omega_o^2\int_0^\infty S_Q(f)\,df. \tag{6.75d}$$

By inserting the spectral density (6.75b) and performing the frequency integral with the help of the narrowness of the resonance, we obtain

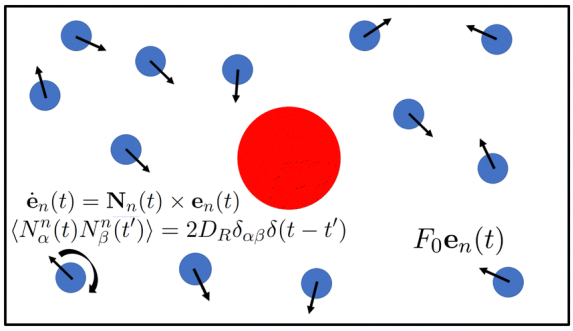
$$\bar{E} = \frac{S_{F'}(f_o)}{4R}. \tag{6.75e}$$

Equating this to our statistical-equilibrium expression (6.75c) for the mean energy, we see that at the frequency $f_o = \omega_o/(2\pi)$ the spectral density $S_{F'}(f_o)$ has the form (6.74b) claimed in the fluctuation-dissipation theorem. Moreover, since f_o can be chosen to be any frequency in the range (6.74a), the spectral density $S_{F'}(f)$ has the claimed form anywhere in this range. ■

218

218

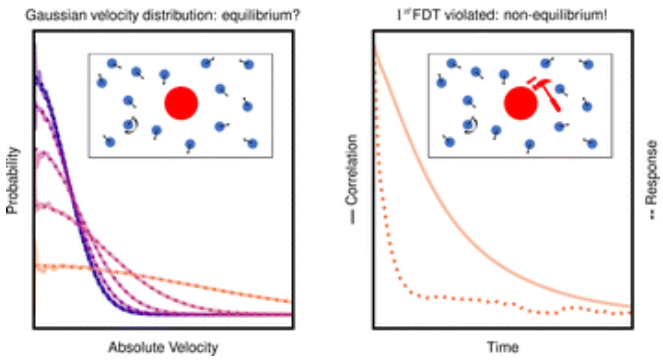
System of a passive probe immersed in a bath of active Langevin particles



219

219

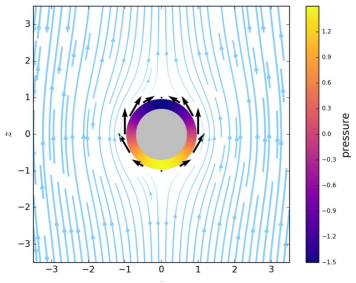
System of a passive probe immersed in a bath of active Langevin particles



Soft Matter, 2022,**18**, 6965-6973

220

220



BROWNIAN MOTION

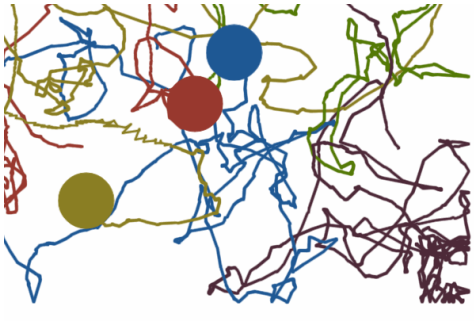
In Secs. 6.3.3 and 6.7.2, we have studied the Brownian motion of a dust particle being buffeted by air molecules, but until now we omitted any attempt to deduce the motion's relaxation time τ_r. We now apply the fluctuation-dissipation theorem to deduce τ_r, using a model in which the particle is idealized as a sphere with mass m and radius a that, of course, is far larger than the air molecules.

The equation of motion for the dust particle, when we ignore the molecules' fluctuating forces, is $mdv/dt = -Rv$. Here the resistance (friction) R due to interaction with the molecules has a form that depends on whether the molecules' mean free path λ is small or large compared to the particle. From the kinetic-theory formula $\lambda = 1/(n\sigma_{\rm mol})$, where n is the number density of molecules and $\sigma_{\rm mol}$ is their cross section to scatter off each other (roughly their cross sectional area), we can deduce that for air $\lambda \sim 0.1\,\mu$m. This is tiny compared to a dust particle's radius $a \sim 10$ to 1,000 μm. This means that, when interacting with the dust particle, the air molecules will behave like a fluid. As we shall learn in Chap. 15, the friction for a fluid depends on whether a quantity called the Reynolds number, $\mathrm{Re} = va/\nu$, is small or large compared to unity; here $\nu \sim 10^{-5}\ \mathrm{m^2\ s^{-1}}$ is the kinematic viscosity of air. Inserting numbers, we see that $\mathrm{Re} \sim (v/0.1\ \mathrm{m\ s^{-1}})(a/100\ \mu\mathrm{m})$. The speeds v of dust particles being buffeted by air are far smaller than $0.1\ \mathrm{m\ s^{-1}}$ as anyone who has watched them in a sunbeam knows, or as you can estimate from Eq. (6.53a). Therefore, the Reynolds number is small. From an analysis carried out in Sec. 14.3.2, we learn that in this low-Re fluid regime, the resistance (friction) on our spherical particle with radius a is [Eq. (14.34)]

$$R = 6\pi\rho\nu a, \tag{6.76}$$

where $\rho \sim 1\ \mathrm{kg\ m^{-3}}$ is the density of air. (Notice that this resistance is proportional to the sphere's radius a or circumference; if λ were $\gg a$, then R would be proportional to the sphere's cross sectional area, i.e., to a^2.)

221



When we turn on the molecules' fluctuating force F', the particle's equation of motion becomes $mdv/dt + Rv = F'$. Feeding $e^{i2\pi ft}$ through this equation in place of F', we get the output $v = 1/(R + i2\pi fm)$, whose modulus squared then is the ratio of S_v to $S_{F'}$. In this obviously classical regime, the fluctuation-dissipation theorem states that $S_{F'} = 4Rk_BT$. Therefore, we have

$$S_v = \frac{S_{F'}}{R^2 + (2\pi fm)^2} = \frac{4Rk_BT}{R^2 + (2\pi fm)^2} = \frac{4Rk_BT/m^2}{(2\pi f)^2 + (R/m)^2}. \tag{6.77}$$

By comparing with the S_v that we derived from Doob's theorem, Eq. (6.53b), we can read off the particle's rms velocity (in one dimension, x or y or z), $\sigma_v = \sqrt{k_BT/m}$—which agrees with Eq. (6.53a) as it must—and we can also read off the particle's relaxation time (not to be confused with the bath's relaxation time),

$$\tau_r = m/R = m/(6\pi\rho\nu a). \tag{6.78}$$

If we had tried to derive this relaxation time by analyzing the buffeting of the particle directly, we would have had great difficulty. The fluctuation-dissipation theorem, Doob's theorem, and the fluid-mechanics analysis of friction on a sphere have made the task straightforward.

222

6.9 Fokker-Planck Equation

In statistical physics, we often want to know the collective influence of many degrees of freedom (a bath) on a single (possibly vectorial) degree of freedom q. The bath might or might not be thermalized. The forces it exerts on q might have short range (as in molecular collisions buffeting an air molecule or dust particle) or long range (as in Coulomb forces from many charged particles in a plasma pushing stochastically on an electron that interests us, or gravitational forces from many stars pulling on a single star of interest). There might also be long-range, macroscopic forces that produce anisotropies and/or inhomogeneities (e.g., applied electric or magnetic fields). We might want to compute how the bath's many degrees of freedom influence, for example, the diffusion of a particle as embodied in its degree of freedom q. Or we might want to compute the statistical properties of q for a representative electron in a plasma and from them deduce the plasma's transport coefficients (diffusivity, heat conductivity, and thermal conductivity). Or we might want to know how the gravitational pulls of many stars in the vicinity of a black hole drive the collective evolution of the stars' distribution function.

223

223

6.9.1 Fokker-Planck for a 1-Dimensional Markov Process

For a 1-dimensional Markov process $y(t)$ (e.g., the x component of the velocity of a particle) being driven by a bath (not necessarily thermalized!) with many degrees of freedom, the *Fokker-Planck equation*[11] states

$$\boxed{\frac{\partial}{\partial t}P_2 = -\frac{\partial}{\partial y}[A(y)P_2] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[B(y)P_2].} \tag{6.94}$$

Here $P_2 = P_2(y, t|y_o)$ is to be regarded as a function of the variables y and t with y_o fixed; that is, Eq. (6.94) is to be solved subject to the initial condition

$$\boxed{P_2(y, 0|y_o) = \delta(y - y_o).} \tag{6.95}$$

As we shall see later, this Fokker-Planck equation is a generalized diffusion equation for the probability P_2: as time passes, the probability diffuses away from its initial location, $y = y_o$, spreading gradually out over a wide range of values of y.

224

224

In the Fokker-Planck equation (6.94) the function $A(y)$ produces a motion of the mean away from its initial location, while the function $B(y)$ produces a diffusion of the probability. If one can deduce the evolution of P_2 for very short times by some other method [e.g., in the case of a dust particle being buffeted by air molecules, by solving statistically the Langevin equation $mdv/dt + Rv = F'(t)$], then from that short-time evolution one can compute the functions $A(y)$ and $B(y)$:

$$A(y) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{+\infty} (y' - y) P_2(y', \Delta t | y) dy', \tag{6.96a}$$

$$B(y) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{+\infty} (y' - y)^2 P_2(y', \Delta t | y) dy'. \tag{6.96b}$$

225

225

(These equations can be deduced by reexpressing the limit as an integral of the time derivative $\partial P_2/\partial t$ and then inserting the Fokker-Planck equation and integrating by parts; Ex. 6.19.) Note that the integral (6.96a) for $A(y)$ is the mean change $\overline{\Delta y}$ in the value of y that occurs in time Δt, if at the beginning of Δt (at $t = 0$) the value of the process is precisely y; moreover (since the integral of yP_2 is just equal to y, which is a constant), $A(y)$ is also the rate of change of the mean, $d\bar{y}/dt$. Correspondingly we can write Eq. (6.96a) in the more suggestive form

$$\boxed{A(y) = \lim_{\Delta t \to 0} \left(\frac{\overline{\Delta y}}{\Delta t} \right) = \left(\frac{d\bar{y}}{dt} \right)_{t=0}.} \tag{6.97a}$$

Similarly, the integral (6.96b) for $B(y)$ is the mean-squared change in y, $\overline{(\Delta y)^2}$, if at the beginning of Δt the value of the process is precisely y; and (as one can fairly easily show; Ex. 6.19) it is also the rate of change of the variance $\sigma_y^2 = \int (y' - \bar{y})^2 P_2 dy'$. Correspondingly, Eq. (6.96b) can be written as

$$\boxed{B(y) = \lim_{\Delta t \to 0} \left(\frac{\overline{(\Delta y)^2}}{\Delta t} \right) = \left(\frac{d\sigma_y^2}{dt} \right)_{t=0}.} \tag{6.97b}$$

226

226

It may seem surprising that $\overline{\Delta y}$ and $\overline{(\Delta y)^2}$ can both increase linearly in time for small times [cf. the Δt in the denominators of both Eq. (6.97a) and Eq. (6.97b)], thereby both giving rise to finite functions $A(y)$ and $B(y)$. In fact, this is so: the linear evolution of $\overline{\Delta y}$ at small t corresponds to the motion of the mean (i.e., of the peak of the probability distribution), while the linear evolution of $\overline{(\Delta y)^2}$ corresponds to the diffusive broadening of the probability distribution.

A solution to the one-dimensional Fokker–Planck equation, with both the drift and the diffusion term. In this case the initial condition is a Dirac delta function centered away from zero velocity. Over time the distribution widens due to random impulses.

227

227

DERIVATION OF THE FOKKER-PLANCK EQUATION (6.94)

Because y is Markov, it satisfies the Smoluchowski equation (6.11), which we rewrite here with a slight change of notation:

$$P_2(y, t+\tau|y_o) = \int_{-\infty}^{+\infty} P_2(y-\xi, t|y_o) P_2(y-\xi+\xi, \tau|y-\xi) d\xi. \quad (6.98a)$$

Take τ to be small so only small ξ will contribute to the integral, and expand in a Taylor series in τ on the left-hand side of (6.98a) and in the ξ of $y-\xi$ on the right-hand side:

$$P_2(y, t|y_o) + \sum_{n=1}^{\infty} \frac{1}{n!} \left[\frac{\partial^n}{\partial t^n} P_2(y, t|y_o) \right] \tau^n$$
$$= \int_{-\infty}^{+\infty} P_2(y, t|y_o) P_2(y+\xi, \tau|y) d\xi$$
$$+ \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} (-\xi)^n \frac{\partial^n}{\partial y^n} [P_2(y, t|y_o) P_2(y+\xi, \tau|y)] \, d\xi. \quad (6.98b)$$

DON'T DRINK AND DERIVE:

$$a = b$$
$$a^2 = ab$$
$$2a^2 = a^2 + ab$$
$$2a^2 - 2ab = a^2 - ab$$
$$2a(a-b) = a(a-b)$$
$$2a = a$$
$$2 = 1$$

228

228

In the first integral on the right-hand side the first term is independent of ξ and can be pulled out from under the integral, and the second term then integrates to one; thereby the first integral on the right reduces to $P_2(y, t|y_o)$, which cancels the first term on the left. The result is then

$$\sum_{n=1}^{\infty} \frac{1}{n!}\left[\frac{\partial^n}{\partial t^n} P_2(y, t|y_o)\right]\tau^n = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}\frac{\partial^n}{\partial y^n}\left[P_2(y, t|y_o)\int_{-\infty}^{+\infty} \xi^n P_2(y+\xi, \tau|y)\, d\xi\right]. \tag{6.98c}$$

Divide by τ, take the limit $\tau \to 0$, and set $\xi \equiv y' - y$ to obtain

$$\frac{\partial}{\partial t} P_2(y, t|y_o) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}\frac{\partial^n}{\partial y^n}[M_n(y) P_2(y, t|y_o)], \tag{6.99a}$$

where

$$M_n(y) \equiv \lim_{\Delta t \to 0} \frac{1}{\Delta t}\int_{-\infty}^{+\infty} (y' - y)^n P_2(y', \Delta t|y)\, dy' \tag{6.99b}$$

is the nth moment of the probability distribution P_2 after time Δt. This is a form of the Fokker-Planck equation that has slightly wider validity than Eq. (6.94). Almost always, however, the only nonvanishing functions $M_n(y)$ are $M_1 \equiv A$, which describes the linear motion of the mean, and $M_2 \equiv B$, which describes the linear growth of the variance. Other moments of P_2 grow as higher powers of Δt than the first power, and correspondingly, their M_ns vanish. Thus, almost always[12] (and always, so far as we are concerned), Eq. (6.99a) reduces to the simpler version (6.94) of the Fokker-Planck equation.

229

DON'T DRINK AND DERIVE:

$$a = b$$
$$a^2 = ab$$
$$2a^2 = a^2 + ab$$
$$2a^2 - 2ab = a^2 - ab$$
$$2a(a-b) = a(a-b)$$
$$2a = a$$
$$2 = 1$$

229

TIME-INDEPENDENT FOKKER-PLANCK EQUATION

If, as we assume in this chapter, y is ergodic, then $p_1(y)$ can be deduced as the limit of $P_2(y, t|y_o)$ for arbitrarily large times t. Then (and in general) p_1 can be deduced from the time-independent Fokker-Planck equation:

$$\boxed{-\frac{\partial}{\partial y}[A(y)p_1(y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[B(y)p_1(y)] = 0.} \tag{6.100}$$

230

230

GAUSSIAN-MARKOV PROCESS

For a Gaussian-Markov process, the mathematical form of $P_2(y_2, t|y_1)$ is known from Doob's theorem: Eqs. (6.18). In the notation of those equations, the Fokker-Planck functions A and B are

$$A(y_1) = (d\bar{y}_t/dt)_{t=0} = -(y_1 - \bar{y})/\tau_r, \quad \text{and} \quad B(y_1) = (d\sigma^2_{y_t}/dt)_{t=0} = 2\sigma^2_y/\tau_r.$$

Translating back to the notation of this section, we have

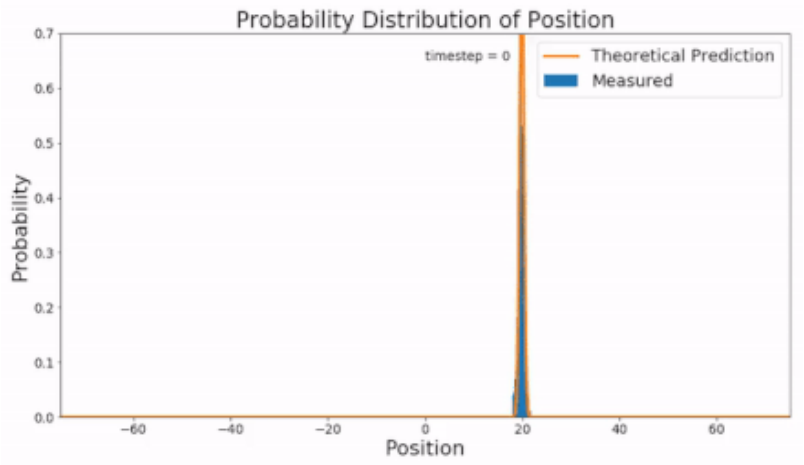
$$A(y) = -(y - \bar{y})/\tau_r, \qquad B(y) = 2\sigma^2_y/\tau_r. \tag{6.101}$$

Thus, if we can compute $A(y)$ and $B(y)$ explicitly for a Gaussian-Markov process, then from them we can read off the process's relaxation time τ_r, long-time mean $\bar{y}$, and long-time variance σ^2_y. As examples, in Ex. 6.22 we revisit Brownian motion of a dust particle in air and in the next section, we analyze laser cooling of atoms. A rather different example is the evolution of a photon distribution function under Compton scattering (Sec. 28.6.3).

231

231

Brownian dynamics simulation for particles in 1-D linear potential compared with the solution of the Fokker–Planck equation



232

232

Bibliographic Note

Random processes are treated in many standard textbooks on statistical physics, typically under the rubric of fluctuations or nonequilibrium statistical mechanics (and sometimes not even using the phrase "random process"). We like Kittel (2004), Sethna (2006), Reif (2008), and Pathria and Beale (2011). A treatise on signal processing that we recommend, despite its age, is Wainstein and Zubakov (1962). There are a number of textbooks on random processes (also called "stochastic processes" in book titles), usually aimed at mathematicians, engineers, or finance folks (who use the theory of random processes to try to make lots of money, and often succeed). But we do not like any of those books as well as the relevant sections in the above statistical mechanics texts. Nevertheless, you might want to peruse Lax et al. (2006), Van Kampen (2007), and Paul and Baschnagel (2010).