Contents

2 Stochastic Processes **1**

2.1 References . . . 1

2.2 Introduction to Stochastic Processes . . . 2

2.2.1 Diffusion and Brownian motion . . . 2

2.2.2 Langevin equation . . . 3

2.3 Distributions and Functionals . . . 6

2.3.1 Basic definitions . . . 6

2.3.2 Correlations for the Langevin equation . . . 8

2.3.3 General ODEs with random forcing . . . 10

2.4 The Fokker-Planck Equation . . . 11

2.4.1 Basic derivation . . . 11

2.4.2 Brownian motion redux . . . 13

2.4.3 Ornstein-Uhlenbeck process . . . 14

2.5 The Master Equation . . . 15

2.5.1 Equilibrium distribution and detailed balance . . . 16

2.5.2 Boltzmann's H-theorem . . . 16

2.5.3 Formal solution to the Master equation . . . 17

2.6 Formal Theory of Stochastic Processes . . . 20

2.6.1 Markov processes . . . 20

2.6.2 Martingales . . . 22

2.6.3 Differential Chapman-Kolmogorov equations . . . 23

2.6.4 Stationary Markov processes and ergodic properties . . . 25

2.6.5 Approach to stationary solution . . . 26

2.7 Appendix : Nonlinear diffusion 27
2.7.1 PDEs with infinite propagation speed 27
2.7.2 The porous medium and p-Laplacian equations 29
2.7.3 Illustrative solutions 30
2.8 Appendix : Langevin equation for a particle in a harmonic well 32
2.9 Appendix : General Linear Autonomous Inhomogeneous ODEs 33
2.9.1 Solution by Fourier transform 33
2.9.2 Higher order ODEs 35
2.9.3 Kramers-Krönig relations 38
2.10 Appendix : Method of Characteristics 40
2.10.1 Quasilinear partial differential equations 40
2.10.2 Example 40

Chapter 2

Stochastic Processes

2.1 References

– C. Gardiner, *Stochastic Methods* (4^{th} edition, Springer-Verlag, 2010)
Very clear and complete text on stochastic methods, with many applications.

– N. G. Van Kampen *Stochastic Processes in Physics and Chemistry* (3^{rd} edition, North-Holland, 2007)
Another standard text. Very readable, but less comprehensive than Gardiner.

– Z. Schuss, *Theory and Applications of Stochastic Processes* (Springer-Verlag, 2010)
In-depth discussion of continuous path stochastic processes and connections to partial differential equations.

– R. Mahnke, J. Kaupužs, and I. Lubashevsky, *Physics of Stochastic Processes* (Wiley, 2009)
Introductory sections are sometimes overly formal, but a good selection of topics.

– A. N. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea, 1956)
The *Urtext* of mathematical probability theory.

2.2 Introduction to Stochastic Processes

A *stochastic process* is one which is partially random, *i.e.* it is not wholly deterministic. Typically the randomness is due to phenomena at the microscale, such as the effect of fluid molecules on a small particle, such as a piece of dust in the air. The resulting motion (called *Brownian motion* in the case of particles moving in a fluid) can be described only in a statistical sense. That is, the full motion of the system is a *functional* of one or more independent random variables. The motion is then described by its averages with respect to the various random distributions.

2.2.1 Diffusion and Brownian motion

Fick's law (1855) is a phenomenological relationship between number current $\boldsymbol{j}$ and number density gradient $\boldsymbol{\nabla} n$, given by $\boldsymbol{j} = -D\,\boldsymbol{\nabla} n$. Combining this with the continuity equation $\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{j}$, one arrives at the diffusion equation[1],

$$\frac{\partial n}{\partial t} = \boldsymbol{\nabla}\cdot(D\,\boldsymbol{\nabla} n)\ . \tag{2.1}$$

Note that the diffusion constant D may be position-dependent. The applicability of Fick's law was experimentally verified in many different contexts and has applicability to a wide range of transport phenomena in physics, chemistry, biology, ecology, geology, *etc.*

The eponymous Robert Brown, a botanist, reported in 1827 on the random motions of pollen grains suspended in water, which he viewed through a microscope. Apparently this phenomenon attracted little attention until the work of Einstein (1905) and Smoluchowski (1906), who showed how it is described by kinetic theory, in which the notion of randomness is essential, and also connecting it to Fick's laws of diffusion. Einstein began with the ideal gas law for osmotic pressure, $p = nk_{\rm B}T$. In steady state, the osmotic force per unit volume acting on the solute (*e.g.* pollen in water), $-\boldsymbol{\nabla} p$, must be balanced by viscous forces. Assuming the solute consists of spherical particles of radius a, the viscous force per unit volume is given by the hydrodynamic *Stokes drag* per particle $\boldsymbol{F} = -6\pi\eta a\boldsymbol{v}$ times the number density n, where η is the dynamical viscosity of the solvent. Thus, $\boldsymbol{j} = n\boldsymbol{v} = -D\,\boldsymbol{\nabla} n$, where $D = k_{\rm B}T/6\pi a\eta$.

To connect this to kinetic theory, Einstein reasoned that the solute particles were being buffeted about randomly by the solvent, and he treated this problem statistically. While a given pollen grain is not significantly effected by any single collision with a water molecule, after some characteristic microscopic time τ the grain has effectively forgotten it initial conditions. Assuming there are no global currents, on average each grain's velocity is zero. Einstein posited that over an interval τ, the number of grains which move a distance within $d^3\Delta$ of $\boldsymbol{\Delta}$ is $n\,\phi(\boldsymbol{\Delta})\,d^3\Delta$, where $\phi(\boldsymbol{\Delta}) = \phi\big(|\boldsymbol{\Delta}|\big)$ is isotropic and also normalized according to $\int d^3\Delta\ \phi(\Delta) = 1$. Then

$$n(\boldsymbol{x}, t+\tau) = \int d^3\Delta\ n(\boldsymbol{x}-\boldsymbol{\Delta}, t)\,\phi(\boldsymbol{\Delta})\ , \tag{2.2}$$

Taylor expanding in both space and time, to lowest order in τ one recovers the diffusion equation, $\partial_t n = D\,\nabla^2 n$, where the diffusion constant is given by

$$D = \frac{1}{6\tau}\int d^3\Delta\ \phi(\Delta)\,\Delta^2\ . \tag{2.3}$$

The diffusion equation with constant D is easily solved by taking the spatial Fourier transform. One then has, in d spatial dimensions,

$$\frac{\partial\hat{n}(\boldsymbol{k},t)}{\partial t} = -D\boldsymbol{k}^2\hat{n}(\boldsymbol{k},t) \qquad \Rightarrow \qquad n(\boldsymbol{x},t) = \int\frac{d^dk}{(2\pi)^d}\,\hat{n}(\boldsymbol{k},t_0)\,e^{-D\boldsymbol{k}^2(t-t_0)}\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\ . \tag{2.4}$$

[1] The equation $\boldsymbol{j} = -D\,\boldsymbol{\nabla} n$ is sometimes called Fick's first law, and the continuity equation $\partial_t n = -\boldsymbol{\nabla}\cdot\boldsymbol{j}$ Fick's second law.

If $n(\boldsymbol{x}, t_0) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$, corresponding to $\hat{n}(\boldsymbol{k}, t_0) = e^{-i\boldsymbol{k}\cdot\boldsymbol{x}_0}$, we have

$$n(\boldsymbol{x}, t) = \big(4\pi D|t - t_0|\big)^{-d/2} \exp\left\{ -\frac{(\boldsymbol{x} - \boldsymbol{x}_0)^2}{4D|t - t_0|} \right\} , \tag{2.5}$$

where d is the dimension of space.

WTF just happened?

We're so used to diffusion processes that most of us overlook a rather striking aspect of the above solution to the diffusion equation. At $t = t_0$, the probability density is $P(\boldsymbol{x}, t = t_0) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$, which means all the particles are sitting at $\boldsymbol{x} = \boldsymbol{x}_0$. For any $t > t_0$, the solution is given by Eqn. 2.5, which is nonzero for all $\boldsymbol{x}$. If we take a value of $\boldsymbol{x}$ such that $|\boldsymbol{x} - \boldsymbol{x}_0| > ct$, where c is the speed of light, we see that there is a finite probability, however small, for particles to diffuse at superluminal speeds. Clearly this is nonsense. The error lies in the diffusion equation itself, which does not recognize any limiting propagation speed. For most processes, this defect is harmless, as we are not interested in the extreme tails of the distribution. Diffusion phenomena and the applicability of the diffusion equation are well-established in virtually every branch of science. To account for a finite propagation speed, one is forced to consider various generalizations of the diffusion equation. Some examples are discussed in the appendix §2.7.

2.2.2 Langevin equation

Consider a particle of mass M subjected to dissipative and random forcing. We'll examine this system in one dimension to gain an understanding of the essential physics. We write

$$\dot{u} + \gamma u = \frac{F}{M} + \eta(t) \, . \tag{2.6}$$

Here, u is the particle's velocity, γ is the damping rate due to friction, F is a constant external force, and $\eta(t)$ is a *stochastic random force*. This equation, known as the Langevin equation, describes a ballistic particle being buffeted by random forcing events[2]. Think of a particle of dust as it moves in the atmosphere. F would then represent the external force due to gravity and $\eta(t)$ the random forcing due to interaction with the air molecules. For a sphere of radius a moving in a fluid of dynamical viscosity η, hydrodynamics gives $\gamma = 6\pi\eta a/M$, where M is the mass of the particle. It is illustrative to compute γ in some setting. Consider a micron sized droplet ($a = 10^{-4}\,\mathrm{cm}$) of some liquid of density $\rho \sim 1.0\,\mathrm{g/cm^3}$ moving in air at $T = 20^\circ\,\mathrm{C}$. The viscosity of air is $\eta = 1.8 \times 10^{-4}\,\mathrm{g/cm \cdot s}$ at this temperature[3]. If the droplet density is constant, then $\gamma = 9\eta/2\rho a^2 = 8.1 \times 10^4\,\mathrm{s^{-1}}$, hence the time scale for viscous relaxation of the particle is $\tau = \gamma^{-1} = 12\,\mu\mathrm{s}$. We should stress that the viscous damping on the particle is of course due to the fluid molecules, in some average 'coarse-grained' sense. The random component to the force $\eta(t)$ would then represent the fluctuations with respect to this average.

We can easily integrate this equation:

$$\begin{aligned} \frac{d}{dt}\left(u\, e^{\gamma t}\right) &= \frac{F}{M}\, e^{\gamma t} + \eta(t)\, e^{\gamma t} \\ u(t) &= u(0)\, e^{-\gamma t} + \frac{F}{\gamma M}\left(1 - e^{-\gamma t}\right) + \int_0^t ds\, \eta(s)\, e^{\gamma(s-t)} \end{aligned} \tag{2.7}$$

[2]See the appendix in §2.8 for the solution of the Langevin equation for a particle in a harmonic well.
[3]The cgs unit of viscosity is the *Poise* (P). $1\,\mathrm{P} = 1\,\mathrm{g/cm \cdot s}$.

Note that $u(t)$ is indeed a functional of the random function $\eta(t)$. We can therefore only compute averages in order to describe the motion of the system.

The first average we will compute is that of v itself. In so doing, we assume that $\eta(t)$ has zero mean: $\big\langle \eta(t) \big\rangle = 0$. Then

$$\big\langle u(t) \big\rangle = u(0)\, e^{-\gamma t} + \frac{F}{\gamma M}\left(1 - e^{-\gamma t}\right) . \tag{2.8}$$

On the time scale γ^{-1}, the initial conditions $u(0)$ are effectively forgotten, and asymptotically for $t \gg \gamma^{-1}$ we have $\big\langle u(t) \big\rangle \to F/\gamma M$, which is the terminal momentum.

Next, consider

$$\big\langle u^2(t) \big\rangle = \big\langle u(t) \big\rangle^2 + \int\limits_0^t ds_1 \int\limits_0^t ds_2 \, e^{\gamma(s_1 - t)} \, e^{\gamma(s_2 - t)} \, \big\langle \eta(s_1)\, \eta(s_2) \big\rangle \ . \tag{2.9}$$

We now need to know the two-time correlator $\big\langle \eta(s_1)\, \eta(s_2) \big\rangle$. We assume that the correlator is a function only of the time difference $\Delta s = s_1 - s_2$, and that the random force $\eta(s)$ has zero average, $\big\langle \eta(s) \big\rangle = 0$, and autocorrelation

$$\big\langle \eta(s_1)\, \eta(s_2) \big\rangle = \phi(s_1 - s_2) \ . \tag{2.10}$$

The function $\phi(s)$ is the *autocorrelation function* of the random force. A macroscopic object moving in a fluid is constantly buffeted by fluid particles over its entire perimeter. These different fluid particles are almost completely uncorrelated, hence $\phi(s)$ is basically nonzero except on a very small time scale τ_ϕ, which is the time a single fluid particle spends interacting with the object. We can take $\tau_\phi \to 0$ and approximate

$$\phi(s) \approx \Gamma\, \delta(s) \ . \tag{2.11}$$

We shall determine the value of Γ from equilibrium thermodynamic considerations below.

With this form for $\phi(s)$, we can easily calculate the equal time momentum autocorrelation:

$$\begin{aligned} \big\langle u^2(t) \big\rangle &= \big\langle u(t) \big\rangle^2 + \Gamma \int\limits_0^t ds \, e^{2\gamma(s-t)} \\ &= \big\langle u(t) \big\rangle^2 + \frac{\Gamma}{2\gamma}\left(1 - e^{-2\gamma t}\right) . \end{aligned} \tag{2.12}$$

Consider the case where $F = 0$ and the limit $t \gg \gamma^{-1}$. We demand that the object thermalize at temperature T. Thus, we impose the condition

$$\big\langle \, \tfrac{1}{2} M u^2(t) \, \big\rangle = \tfrac{1}{2} k_{\rm B} T \qquad \Longrightarrow \qquad \Gamma = \frac{2\gamma k_{\rm B} T}{M} \quad . \tag{2.13}$$

This fixes the value of Γ.

We can now compute the general momentum autocorrelator:

$$\begin{aligned} \big\langle u(t)\, u(t') \big\rangle - \big\langle u(t) \big\rangle \big\langle u(t') \big\rangle &= \int\limits_0^t ds \int\limits_0^{t'} ds' \, e^{\gamma(s-t)} \, e^{\gamma(s'-t')} \, \big\langle \eta(s)\, \eta(s') \big\rangle \\ &= \frac{\Gamma}{2\gamma}\, e^{-\gamma|t-t'|} \qquad (t, t' \to \infty \ , \ |t - t'| \ \text{finite}) \ . \end{aligned} \tag{2.14}$$

Let's now compute the position $x(t)$. We find

$$x(t) = \big\langle x(t) \big\rangle + \frac{1}{M} \int\limits_0^t ds \int\limits_0^s ds_1 \, \eta(s_1) \, e^{\gamma(s_1 - s)} \ , \tag{2.15}$$

where

$$\big\langle x(t) \big\rangle = x(0) + \frac{1}{\gamma} \left(u(0) - \frac{F}{\gamma M} \right) \left(1 - e^{-\gamma t} \right) + \frac{Ft}{\gamma M} \ . \tag{2.16}$$

Note that for $\gamma t \ll 1$ we have $\big\langle x(t) \big\rangle = x(0) + u(0)\, t + \frac{1}{2} M^{-1} F t^2 + \mathcal{O}(t^3)$, as is appropriate for ballistic particles moving under the influence of a constant force. This long time limit of course agrees with our earlier evaluation for the terminal velocity, $\big\langle u(\infty) \big\rangle = F/\gamma M$. We next compute the position autocorrelation:

$$\begin{aligned} \big\langle x(t)\, x(t') \big\rangle - \big\langle x(t) \big\rangle \big\langle x(t') \big\rangle &= \frac{1}{M^2} \int\limits_0^t ds \int\limits_0^{t'} ds' \, e^{-\gamma(s+s')} \int\limits_0^s ds_1 \int\limits_0^{s'} ds_1' \, e^{\gamma(s_1 + s_2)} \, \big\langle \eta(s_1)\, \eta(s_2) \big\rangle \\ &= \frac{2 k_{\rm B} T}{\gamma M} \mathsf{min}(t, t') + \mathcal{O}(1) \ . \end{aligned}$$

In particular, the equal time autocorrelator is

$$\big\langle x^2(t) \big\rangle - \big\langle x(t) \big\rangle^2 = \frac{2 k_{\rm B} T \, t}{\gamma M} \equiv 2D\, t \ , \tag{2.17}$$

at long times, up to terms of order unity. Here, $D = \Gamma/2\gamma^2 = k_{\rm B} T/\gamma M$ is the *diffusion constant*. For a liquid droplet of radius $a = 1\ \mu{\rm m}$ moving in air at $T = 293\,{\rm K}$, for which $\eta = 1.8 \times 10^{-4}\,{\rm P}$, we have

$$D = \frac{k_{\rm B} T}{6\pi\eta a} = \frac{(1.38 \times 10^{-16}\,{\rm erg/K})\,(293\,{\rm K})}{6\pi\,(1.8 \times 10^{-4}\,{\rm P})\,(10^{-4}\,{\rm cm})} = 1.19 \times 10^{-7}\,{\rm cm}^2/{\rm s} \ . \tag{2.18}$$

This result presumes that the droplet is large enough compared to the intermolecular distance in the fluid that one can adopt a continuum approach and use the Navier-Stokes equations, and then assuming a laminar flow.

If we consider molecular diffusion, the situation is quite a bit different. The diffusion constant is then $D = \ell^2/2\tau$, where ℓ is the mean free path and τ is the collision time. Elementary kinetic theory gives that the mean free path ℓ, collision time τ, number density n, and total scattering cross section σ are related by[4] $\ell = \bar{v}\tau = 1/\sqrt{2}\, n\sigma$, where $\bar{v} = \sqrt{8 k_{\rm B} T/\pi m}$ is the average particle speed. Approximating the particles as hard spheres, we have $\sigma = 4\pi a^2$, where a is the hard sphere radius. At $T = 293\,{\rm K}$, and $p = 1\,{\rm atm}$, we have $n = p/k_{\rm B} T = 2.51 \times 10^{19}\,{\rm cm}^{-3}$. Since air is predominantly composed of ${\rm N}_2$ molecules, we take $a = 1.90 \times 10^{-8}\,{\rm cm}$ and $m = 28.0\,{\rm amu} = 4.65 \times 10^{-23}\,{\rm g}$, which are appropriate for ${\rm N}_2$. We find an average speed of $\bar{v} = 471\,{\rm m/s}$ and a mean free path of $\ell = 6.21 \times 10^{-6}\,{\rm cm}$. Thus, $D = \frac{1}{2}\ell\bar{v} = 0.146\,{\rm cm}^2/{\rm s}$. Though much larger than the diffusion constant for large droplets, this is still too small to explain common experiences. Suppose we set the characteristic distance scale at $d = 10\,{\rm cm}$ and we ask how much time a point source would take to diffuse out to this radius. The answer is $\Delta t = d^2/2D = 343\,{\rm s}$, which is between five and six minutes. Yet if someone in the next seat emits a foul odor, you detect the offending emission in on the order of a second. What this tells us is that diffusion isn't the only transport process involved in these and like phenomena. More important are *convection* currents which distribute the scent much more rapidly.

[4] The scattering time τ is related to the particle density n, total scattering cross section σ, and mean speed $\bar{v}$ through the relation $n\sigma\bar{v}_{\rm rel}\tau = 1$, which says that on average one scattering event occurs in a cylinder of cross section σ and length $\bar{v}_{\rm rel}\tau$. Here $\bar{v}_{\rm rel} = \sqrt{\bar{v}}$ is the mean relative speed of a pair of particles.

2.3 Distributions and Functionals

2.3.1 Basic definitions

Let $x \in \mathbb{R}$ be a random variable, and $P(x)$ a probability distribution for x. The average of any function $\phi(x)$ is then

$$\big\langle \phi(x) \big\rangle = \int\limits_{-\infty}^{\infty} dx\ P(x)\,\phi(x) \Big/ \int\limits_{-\infty}^{\infty} dx\ P(x)\ . \tag{2.19}$$

Let $\eta(t)$ be a random *function* of t, with $\eta(t) \in \mathbb{R}$, and let $P\big[\eta(t)\big]$ be the probability distribution *functional* for $\eta(t)$. Then if $\Phi\big[\eta(t)\big]$ is a functional of $\eta(t)$, the average of Φ is given by

$$\int D\eta\ P\big[\eta(t)\big]\ \Phi\big[\eta(t)\big] \Big/ \int D\eta\ P\big[\eta(t)\big] \tag{2.20}$$

The expression $\int D\eta\ P[\eta]\,\Phi[\eta]$ is a *functional integral*. A functional integral is a continuum limit of a multivariable integral. Suppose $\eta(t)$ were defined on a set of t values $t_n = n\tau$. A functional of $\eta(t)$ becomes a multivariable function of the values $\eta_n \equiv \eta(t_n)$. The metric then becomes $D\eta = \prod_n d\eta_n$.

In fact, for our purposes we will not need to know any details about the functional measure $D\eta$; we will finesse this delicate issue[5]. Consider the *generating functional*,

$$Z\big[J(t)\big] = \int D\eta\ P[\eta]\ \exp\left\{ \int\limits_{-\infty}^{\infty} dt\ J(t)\,\eta(t) \right\}\ . \tag{2.21}$$

It is clear that

$$\frac{1}{Z[J]}\ \frac{\delta^n Z[J]}{\delta J(t_1)\cdots\delta J(t_n)}\bigg|_{J(t)=0} = \big\langle \eta(t_1)\cdots\eta(t_n) \big\rangle\ . \tag{2.22}$$

The function $J(t)$ is an arbitrary *source function*. We functionally differentiate with respect to it in order to find the η-field correlators. The functional derivative $\delta Z\big[J(t)\big]/\delta J(s)$ can be computed by substituting $J(t) \to J(t) + \epsilon\,\delta(t-s)$ inside the functional $Z[J]$, and then taking the ordinary derivative with respect to ε, *i.e.*

$$\frac{\delta Z\big[J(t)\big]}{\delta J(s)} = \frac{dZ\big[J(t) + \varepsilon\,\delta(t-s)\big]}{d\varepsilon}\bigg|_{\varepsilon=0}\ . \tag{2.23}$$

Thus the functional derivative $\delta Z\big[J(t)\big]/\delta J(s)$ tells us how the functional $Z[J]$ changes when the function $J(t)$ is replaced by $J(t) + \varepsilon\,\delta(t-s)$. Equivalently, one may eschew this ε prescription and use the familiar chain rule from differential calculus, supplemented by the rule $\delta J(t)/\delta J(s) = \delta(t-s)$.

Let's compute the generating functional for a class of distributions of the Gaussian form,

$$\begin{aligned} P[\eta] &= \exp\left\{ -\frac{1}{2\Gamma} \int\limits_{-\infty}^{\infty} dt\ \left(\tau^2\,\dot\eta^2 + \eta^2\right) \right\} \\ &= \exp\left\{ -\frac{1}{2\Gamma} \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\ \left(1 + \omega^2\tau^2\right) \big|\hat\eta(\omega)\big|^2 \right\}\ . \end{aligned} \tag{2.24}$$

[5] A discussion of measure for functional integrals is found in R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*.

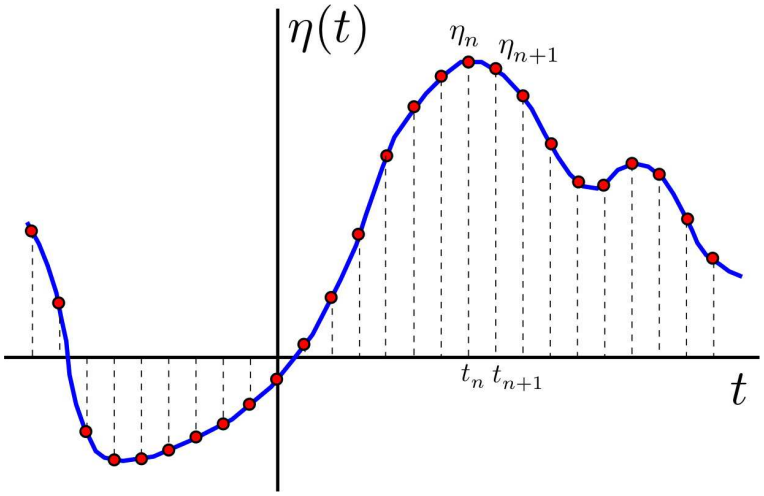

Figure 2.1: Discretization of a continuous function $\eta(t)$. Upon discretization, a functional $\Phi\big[\eta(t)\big]$ becomes an ordinary multivariable function $\Phi(\{\eta_j\})$.

Then Fourier transforming the source function $J(t)$, it is easy to see that

$$Z[J] = Z[0] \cdot \exp\left\{ \frac{\Gamma}{2} \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi} \, \frac{\big|\hat{J}(\omega)\big|^2}{1+\omega^2\tau^2} \right\} . \tag{2.25}$$

Note that with $\eta(t) \in \mathbb{R}$ and $J(t) \in \mathbb{R}$ we have $\eta^*(\omega) = \eta(-\omega)$ and $\hat{J}^*(\omega) = \hat{J}(-\omega)$. Transforming back to real time, we have

$$Z[J] = Z[0] \cdot \exp\left\{ \frac{1}{2} \int\limits_{-\infty}^{\infty} dt \int\limits_{-\infty}^{\infty} dt' \; J(t)\, G(t-t')\, J(t') \right\} , \tag{2.26}$$

where

$$G(s) = \frac{\Gamma}{2\tau}\, e^{-|s|/\tau} \qquad , \qquad \widehat{G}(\omega) = \frac{\Gamma}{1+\omega^2\tau^2} \tag{2.27}$$

is the *Green's function*, in real and Fourier space. Note that

$$\int\limits_{-\infty}^{\infty} ds \, G(s) = \widehat{G}(0) = \Gamma \, . \tag{2.28}$$

We can now compute

$$\big\langle \eta(t_1)\, \eta(t_2) \big\rangle = G(t_1 - t_2) \tag{2.29}$$

$$\begin{aligned} \big\langle \eta(t_1)\, \eta(t_2)\, \eta(t_3)\, \eta(t_4) \big\rangle = {} & G(t_1-t_2)\, G(t_3-t_4) + G(t_1-t_3)\, G(t_2-t_4) \\ & + G(t_1-t_4)\, G(t_2-t_3) \, . \end{aligned} \tag{2.30}$$

The generalization is now easy to prove, and is known as *Wick's theorem*:

$$\big\langle \eta(t_1) \cdots \eta(t_{2n}) \big\rangle = \sum_{\text{contractions}} G(t_{i_1} - t_{i_2}) \cdots G(t_{i_{2n-1}} - t_{i_{2n}}) \, , \tag{2.31}$$

where the sum is over all distinct *contractions* of the sequence $1\cdot 2\,\cdots\,2n$ into products of pairs. How many terms are there? Some simple combinatorics answers this question. Choose the index 1. There are $(2n-1)$ other time indices with which it can be contracted. Now choose another index. There are $(2n-3)$ indices with which *that* index can be contracted. And so on. We thus obtain

$$C(n) \equiv \left\{ \begin{matrix} \text{\# of contractions} \\ \text{of 1-2-3}\cdots 2n \end{matrix} \right\} = (2n-1)(2n-3)\,\cdots\,3\cdot 1 = \frac{(2n)!}{2^n\,n!}\,. \tag{2.32}$$

2.3.2 Correlations for the Langevin equation

Now suppose we have the Langevin equation

$$\frac{du}{dt} + \gamma u = \eta(t) \tag{2.33}$$

with $u(0)=0$. We wish to compute the joint probability density

$$P(u_1,t_1;\ldots;u_N,t_N) = \Big\langle \delta\big(u_1 - u(t_1)\big)\cdots\delta\big(u_N - u(t_N)\big)\Big\rangle\,, \tag{2.34}$$

where the average is over all realizations of the random variable $\eta(t)$:

$$\big\langle F\big[\eta(t)\big]\big\rangle = \int D\eta\, P\big[\eta(t)\big]\, F\big[\eta(t)\big]\,. \tag{2.35}$$

Using the integral representation of the Dirac δ-function, we have

$$P(u_1,t_1;\ldots;u_N,t_N) = \int\limits_0^\infty \frac{d\omega_1}{2\pi}\,\cdots\int\limits_0^\infty \frac{d\omega_N}{2\pi}\,e^{-i(\omega_1 u_1+\ldots+\omega_N u_N)}\,\Big\langle e^{i\omega_1 u(t_1)}\cdots e^{i\omega_N u(t_N)}\Big\rangle\,. \tag{2.36}$$

Now integrating the Langevin equation with the initial condition $u(0)=0$ gives

$$u(t_j) = \int\limits_0^{t_j} dt\, e^{\gamma(t-t_j)}\,\eta(t)\,, \tag{2.37}$$

and therefore we may write

$$\sum_{j=1}^N \omega_j\, u(t_j) = \int\limits_{-\infty}^{\infty} dt\, f(t)\,\eta(t) \tag{2.38}$$

with

$$f(t) = \sum_{j=1}^N \omega_j\, e^{\gamma(t-t_j)}\,\Theta(t)\,\Theta(t_j-t)\,. \tag{2.39}$$

We assume that the random variable $\eta(t)$ is distributed as a Gaussian, with $\big\langle \eta(t)\,\eta(t')\big\rangle = G(t-t')$, as described above. Using our previous results, we may perform the functional integral over $\eta(t)$ to obtain

$$\begin{aligned}\Big\langle \exp\, i\int\limits_{-\infty}^{\infty} dt\, f(t)\,\eta(t)\Big\rangle &= \exp\left\{-\frac{1}{2}\int\limits_{-\infty}^{\infty}dt\int\limits_{-\infty}^{\infty}dt'\,G(t-t')\,f(t)\,f(t')\right\} \\ &= \exp\left\{-\frac{1}{2}\sum_{j,j'=1}^N M_{jj'}\,\omega_j\,\omega_{j'}\right\},\end{aligned} \tag{2.40}$$

where $M_{jj'} = M(t_j, t_{j'})$ with

$$M(t,t') = \int_0^t ds \int_0^{t'} ds'\, G(s-s')\, e^{\gamma(s-t)}\, e^{\gamma(s'-t')}\ . \tag{2.41}$$

We now have

$$\begin{aligned} P(u_1,t_1;\ldots;u_N,t_N) &= \int_0^\infty \frac{d\omega_1}{2\pi}\ \cdots \int_0^\infty \frac{d\omega_N}{2\pi}\ e^{-i(\omega_1 u_1 + \ldots + \omega_N u_N)}\ \exp\left\{ -\tfrac{1}{2} \sum_{j,j'=1}^N M_{jj'}\, \omega_j\, \omega_{j'} \right\} \\ &= \det{}^{-1/2}(2\pi M)\ \exp\left\{ -\tfrac{1}{2} \sum_{j,j'=1}^N M^{-1}_{jj'}\, u_j\, u_{j'} \right\} . \end{aligned} \tag{2.42}$$

In the limit $G(s) = \Gamma\, \delta(s)$, we have

$$\begin{aligned} M_{jj'} &= \Gamma \int_0^{\min(t_j,t_{j'})} dt\, e^{2\gamma t}\, e^{-\gamma(t_j+t_{j'})} \\ &= \frac{\Gamma}{2\gamma} \left(e^{-\gamma|t_j - t_{j'}|} - e^{-\gamma(t_j+t_{j'})} \right) . \end{aligned} \tag{2.43}$$

From this and the previous expression, we have, assuming $t_{1,2} \gg \gamma^{-1}$ but making no assumptions about the size of $|t_1 - t_2|$,

$$P(u_1,t_1) = \sqrt{\frac{\gamma}{\pi\Gamma}}\, e^{-\gamma u_1^2/\Gamma}\ . \tag{2.44}$$

The conditional distribution $P(u_1,t_1\,|\,u_2,t_2) = P(u_1,t_1;u_2,t_2)/P(u_2,t_2)$ is found to be

$$P(u_1,t_1\,|\,u_2,t_2) = \sqrt{\frac{\gamma/\pi\Gamma}{1-e^{-2\gamma(t_1-t_2)}}}\ \exp\left\{ -\frac{\gamma}{\Gamma} \cdot \frac{\left(u_1 - e^{-\gamma(t_1-t_2)}\, u_2\right)^2}{1-e^{-2\gamma(t_1-t_2)}} \right\}\ . \tag{2.45}$$

Note that $P(u_1,t_1\,|\,u_2,t_2)$ tends to $P(u_1,t_1)$ independent of the most recent condition, in the limit $t_1 - t_2 \gg \gamma^{-1}$.

As we shall discuss below, a *Markov process* is one where, at any given time, the statistical properties of the subsequent evolution are fully determined by state of the system at that time. Equivalently, every *conditional probability* depends only on the *most recent condition*. Is $u(t)$ a continuous time Markov process? Yes it is! The reason is that $u(t)$ satisfies a first order differential equation, hence only the initial condition on u is necessary in order to derive its probability distribution at any time in the future. Explicitly, we can compute $P(u_1t_1|\,u_2t_2,u_3t_3)$ and show that it is independent of u_3 and t_3 for $t_1 > t_2 > t_3$. This is true regardless of the relative sizes of $t_j - t_{j+1}$ and γ^{-1}.

While $u(t)$ defines a Markov process, its integral $x(t)$ does not. This is because more information than the initial value of x is necessary in order to integrate forward to a solution at future times. Since $x(t)$ satisfies a second order ODE, its conditional probabilities should in principle depend only on the *two most recent conditions*. We could also consider the evolution of the pair $\boldsymbol{\varphi} = (x,u)$ in *phase space*, writing

$$\frac{d}{dt} \begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & -\gamma \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} + \begin{pmatrix} 0 \\ \eta(t) \end{pmatrix}\ , \tag{2.46}$$

or $\dot{\boldsymbol{\varphi}} = A\boldsymbol{\varphi} + \boldsymbol{\eta}(t)$, where A is the above 2×2 matrix, and the stochastic term $\boldsymbol{\eta}(t)$ has only a lower component. The paths $\boldsymbol{\varphi}(t)$ are also Markovian, because they are determined by a first order set of coupled ODEs. In the limit where $t_j - t_{j+1} \gg \gamma^{-1}$, $x(t)$ effectively becomes Markovian, because we interrogate the paths on time scales where the separations are such that the particle has 'forgotten' its initial velocity.

2.3.3 General ODEs with random forcing

Now let's make a leap to the general $n^{\rm th}$ order linear autonomous inhomogeneous ODE

$$\mathcal{L}_t\, x(t) = \eta(t)\ , \tag{2.47}$$

where $\eta(t)$ is a random function and where

$$\mathcal{L}_t = a_n\,\frac{d^n}{dt^n} + a_{n-1}\,\frac{d^{n-1}}{dt^{n-1}} + \cdots + a_1\,\frac{d}{dt} + a_0 \tag{2.48}$$

is an $n^{\rm th}$ order differential operator. We are free, without loss of generality, to choose $a_n = 1$. In the appendix in §2.9 we solve this equation using a Fourier transform method. But if we want to impose a boundary condition at $t = 0$, it is more appropriate to consider a *Laplace* transform.

The Laplace transform $\check{x}(z)$ is obtained from a function $x(t)$ via

$$\check{x}(z) = \int\limits_0^\infty\! dt\; e^{-zt}\, x(t)\ . \tag{2.49}$$

The inverse transform is given by

$$x(t) = \tfrac{1}{2\pi i}\int\limits_{c-i\infty}^{c+i\infty}\! dz\; e^{zt}\,\check{x}(z)\ , \tag{2.50}$$

where the integration contour is a straight line which lies to the right of any singularities of $\check{x}(z)$ in the complex z plane. Now let's take the Laplace transform of Eqn. 2.47. Note that integration by parts yields

$$\int\limits_0^\infty\! dt\; e^{-zt}\,\frac{df}{dt} = z\check{f}(z) - f(0) \tag{2.51}$$

for any function $f(t)$. Applying this result iteratively, we find that the Laplace transform of Eqn. 2.47 is

$$L(z)\,\check{x}(z) = \check{\eta}(z) + R_0(z)\ , \tag{2.52}$$

where

$$L(z) = a_n z^n + a_{n-1}z^{n-1} + \ldots + a_0 \tag{2.53}$$

is an $n^{\rm th}$ order polynomial in z with coefficients a_j for $j \in \{0,\ldots,n\}$, and

$$R_0(z) = a_n\,x^{(n-1)}(0) + \big(za_n + a_{n-1}\big)\,x^{(n-2)}(0) + \cdots + \big(z^{n-1}a_n + \ldots + a_1\big)\,x(0) \tag{2.54}$$

and $x^{(k)}(t) = d^kx/dt^k$. We now have

$$\check{x}(z) = \frac{1}{L(z)}\,\Big\{\check{\eta}(z) + R_0(z)\Big\}\ . \tag{2.55}$$

The formal solution to Eqn. 2.47 is then given by the inverse Laplace transform. One finds

$$x(t) = \int\limits_0^t\! dt'\; K(t-t')\,\eta(t') + x_{\rm h}(t)\ , \tag{2.56}$$

where $x_{\rm h}(t)$ is a solution to the homogeneous equation $\mathcal{L}_t\, x(t) = 0$, and

$$K(s) = \tfrac{1}{2\pi i}\int\limits_{c-i\infty}^{c+i\infty} dz\ \frac{e^{zs}}{L(z)} = \sum_{l=1}^{n} \frac{e^{z_l s}}{L'(z_l)}\ . \tag{2.57}$$

Note that $K(s)$ vanishes for $s < 0$ because then we can close the contour in the far right half plane. The RHS of the above equation follows from the fundamental theorem of algebra, which allows us to factor $L(z)$ as

$$L(z) = a_n(z - z_1)\cdots(z - z_n)\ , \tag{2.58}$$

with all the roots z_l lying to the left of the contour. In deriving the RHS of Eqn. 2.57, we assume that all roots are distinct[6]. The general solution to the homogeneous equation is

$$x_{\rm h}(t) = \sum_{l=1}^{n} A_l\, e^{z_l t}\ , \tag{2.59}$$

again assuming the roots are nondegenerate[7]. In order that the homogeneous solution not grow with time, we must have $\mathsf{Re}\,(z_l) \le 0$ for all l.

For example, if $\mathcal{L}_t = \frac{d}{dt} + \gamma$, then $L(z) = z + \gamma$ and $K(s) = e^{-\gamma s}$. If $\mathcal{L}_t = \frac{d^2}{dt^2} + \gamma\frac{d}{dt}$, then $L(z) = z^2 + \gamma z$ and $K(s) = (1 - e^{-\gamma s})/\gamma$.

Let us assume that all the initial derivatives $d^k x(t)/dt^k$ vanish at $t = 0$, hence $x_{\rm h}(t) = 0$. Now let us compute the generalization of Eqn. 2.36,

$$\begin{aligned} P(x_1, t_1; \ldots; x_N, t_N) &= \int\limits_0^\infty \frac{d\omega_1}{2\pi}\ \cdots \int\limits_0^\infty \frac{d\omega_N}{2\pi}\ e^{-i(\omega_1 x_1 + \ldots + \omega_N x_N)}\ \Big\langle e^{i\omega_1 x(t_1)}\cdots e^{i\omega_N x(t_N)}\Big\rangle \\ &= \mathsf{det}^{-1/2}(2\pi M)\ \exp\left\{ -\tfrac{1}{2}\sum_{j,j'=1}^{N} M^{-1}_{jj'}\, x_j\, x_{j'} \right\}\ , \end{aligned} \tag{2.60}$$

where

$$M(t, t') = \int\limits_0^t ds \int\limits_0^{t'} ds'\ G(s - s')\, K(t - s)\, K(t' - s')\ , \tag{2.61}$$

with $G(s - s') = \big\langle \eta(s)\,\eta(s')\big\rangle$ as before. For $t \gg \gamma^{-1}$, we have $K(s) = \gamma^{-1}$, and if we take $G(s - s') = \Gamma\,\delta(s - s')$ we obtain $M(t, t') = \Gamma\,\mathsf{min}(t, t')/\gamma^2 = 2D\,\mathsf{min}(t, t')$. We then have $P(x, t) = \exp\big(-x^2/4Dt\big)/\sqrt{4\pi D t}$, as expected.

2.4 The Fokker-Planck Equation

2.4.1 Basic derivation

Suppose $x(t)$ is a stochastic variable. We define the quantity

$$\delta x(t) \equiv x(t + \delta t) - x(t)\ , \tag{2.62}$$

[6] If two or more roots are degenerate, one can still use this result by first inserting a small spacing ε between the degenerate roots and then taking $\varepsilon \to 0$.

[7] If a particular root z_j appears k times, then one has solutions of the form $e^{z_j t}$, $t\, e^{z_j t}$, $\ldots$ $t^{k-1}\, e^{z_j t}$.

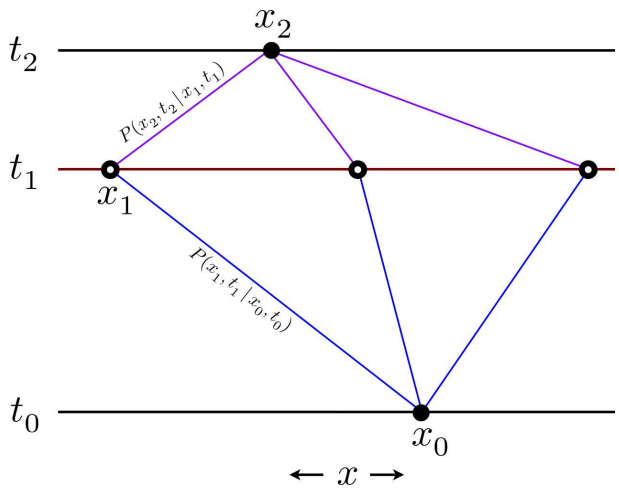

Figure 2.2: Interpretive sketch of the mathematics behind the Chapman-Kolmogorov equation.

and we assume

$$\big\langle \delta x(t) \big\rangle = F_1\big(x(t)\big)\,\delta t \tag{2.63}$$

$$\big\langle \big[\delta x(t)\big]^2 \big\rangle = F_2\big(x(t)\big)\,\delta t \tag{2.64}$$

but $\big\langle \big[\delta x(t)\big]^n \big\rangle = \mathcal{O}\big((\delta t)^2\big)$ for $n > 2$. The $n = 1$ term is due to *drift* and the $n = 2$ term is due to *diffusion*. Now consider the conditional probability density, $P(x, t\,|\,x_0, t_0)$, defined to be the probability distribution for $x \equiv x(t)$ given that $x(t_0) = x_0$. The conditional probability density satisfies the composition rule,

$$P(x_2, t_2\,|\,x_0, t_0) = \int\limits_{-\infty}^{\infty} dx_1\, P(x_2, t_2\,|\,x_1, t_1)\, P(x_1, t_1\,|\,x_0, t_0)\;, \tag{2.65}$$

for any value of t_1. This is also known as the *Chapman-Kolmogorov equation*. In words, what it says is that the probability density for a particle being at x_2 at time t_2, given that it was at x_0 at time t_0, is given by the product of the probability density for being at x_2 at time t_2 given that it was at x_1 at t_1, multiplied by that for being at x_1 at t_1 given it was at x_0 at t_0, integrated over x_1. This should be intuitively obvious, since if we pick any time $t_1 \in [t_0, t_2]$, then the particle had to be *somewhere* at that time. What is perhaps not obvious is why the conditional probability $P(x_2, t_2\,|\,x_1, t_1)$ does not also depend on (x_0, t_0). This is so if the system is described by a *Markov process*, about we shall have more to say below in §2.6.1. At any rate, a picture is worth a thousand words: see Fig. 2.2.

Proceeding, we may write

$$P(x, t+\delta t\,|\,x_0, t_0) = \int\limits_{-\infty}^{\infty} dx'\, P(x, t+\delta t\,|\,x', t)\, P(x', t\,|\,x_0, t_0)\;. \tag{2.66}$$

Now

$$\begin{aligned} P(x, t+\delta t\,|\,x', t) &= \big\langle \delta\big(x - \delta x(t) - x'\big)\big\rangle \\ &= \left\{1 + \big\langle \delta x(t)\big\rangle \frac{d}{dx'} + \tfrac{1}{2}\big\langle \big[\delta x(t)\big]^2\big\rangle \frac{d^2}{dx'^2} + \ldots \right\} \delta(x - x') \\ &= \delta(x - x') + F_1(x')\,\frac{d\,\delta(x-x')}{dx'}\,\delta t + \tfrac{1}{2} F_2(x')\,\frac{d^2\delta(x-x')}{dx'^2}\,\delta t + \mathcal{O}\big((\delta t)^2\big)\;, \end{aligned} \tag{2.67}$$

where the average is over the random variables. We now insert this result into eqn. 2.66, integrate by parts, divide by δt, and then take the limit $\delta t \to 0$. The result is the Fokker-Planck equation,

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}\big[F_1(x)\,P(x,t)\big] + \frac{1}{2}\,\frac{\partial^2}{\partial x^2}\big[F_2(x)\,P(x,t)\big]\ . \tag{2.68}$$

2.4.2 Brownian motion redux

Let's apply our Fokker-Planck equation to a description of Brownian motion. From our earlier results, we have $F_1(x) = F/\gamma M$ and $F_2(x) = 2D$. A formal proof of these results is left as an exercise for the reader. The Fokker-Planck equation is then

$$\frac{\partial P}{\partial t} = -u\,\frac{\partial P}{\partial x} + D\,\frac{\partial^2 P}{\partial x^2}\ , \tag{2.69}$$

where $u = F/\gamma M$ is the average terminal velocity. If we make a Galilean transformation and define $y = x - ut$ and $s = t$, then our Fokker-Planck equation takes the form

$$\frac{\partial P}{\partial s} = D\,\frac{\partial^2 P}{\partial y^2}\ . \tag{2.70}$$

This is known as the *diffusion equation*. Eqn. 2.69 is also a diffusion equation, rendered in a moving frame.

While the Galilean transformation is illuminating, we can easily solve eqn. 2.69 without it. Let's take a look at this equation after Fourier transforming from x to q:

$$P(x,t) = \int\limits_{-\infty}^{\infty}\frac{dq}{2\pi}\,e^{iqx}\,\hat{P}(q,t) \tag{2.71}$$

$$\hat{P}(q,t) = \int\limits_{-\infty}^{\infty}dx\,e^{-iqx}\,P(x,t)\ . \tag{2.72}$$

Then as should be well known to you by now, we can replace the operator $\frac{\partial}{\partial x}$ with multiplication by iq, resulting in

$$\frac{\partial}{\partial t}\,\hat{P}(q,t) = -(Dq^2 + iqu)\,\hat{P}(q,t)\ , \tag{2.73}$$

with solution

$$\hat{P}(q,t) = e^{-Dq^2t}\,e^{-iqut}\,\hat{P}(q,0)\ . \tag{2.74}$$

We now apply the inverse transform to get back to x-space:

$$\begin{aligned} P(x,t) &= \int\limits_{-\infty}^{\infty}\frac{dq}{2\pi}\,e^{iqx}\,e^{-Dq^2t}\,e^{-iqut}\int\limits_{-\infty}^{\infty}dx'\,e^{-iqx'}\,P(x',0) \\ &= \int\limits_{-\infty}^{\infty}dx'\,P(x',0)\int\limits_{-\infty}^{\infty}\frac{dq}{2\pi}\,e^{-Dq^2t}\,e^{iq(x-ut-x')} = \int\limits_{-\infty}^{\infty}dx'\,K(x-x',t)\,P(x',0)\ , \end{aligned} \tag{2.75}$$

where

$$K(x,t) = \frac{1}{\sqrt{4\pi Dt}}\,e^{-(x-ut)^2/4Dt} \tag{2.76}$$

is the *diffusion kernel*. We now have a recipe for obtaining $P(x,t)$ given the initial conditions $P(x,0)$. If $P(x,0) = \delta(x)$, describing a particle confined to an infinitesimal region about the origin, then $P(x,t) = K(x,t)$ is the probability distribution for finding the particle at x at time t. There are two aspects to $K(x,t)$ which merit comment. The first is that the center of the distribution moves with velocity u. This is due to the presence of the external force. The second is that the standard deviation $\sigma = \sqrt{2Dt}$ is increasing in time, so the distribution is not only shifting its center but it is also getting broader as time evolves. This movement of the center and broadening are what we have called *drift* and *diffusion*, respectively.

2.4.3 Ornstein-Uhlenbeck process

Starting from any initial condition $P(x,0)$, the Fokker-Planck equation for Brownian motion, even with drift, inexorably evolves the distribution $P(x,t)$ toward an infinitesimal probability uniformly spread throughout all space. Consider now the Fokker-Planck equation with $F_2(x) = 2D$ as before, but with $F_1(x) = -\beta x$. Thus we have diffusion but also drift, where the local velocity is $-\beta x$. For $x > 0$, probability which diffuses to the right will also drift to the left, so there is a competition between drift and diffusion. Who wins?

We can solve this model exactly. Starting with the FPE

$$\partial_t P = \partial_x(\beta x P) + D\,\partial_x^2 P\ , \tag{2.77}$$

we first Fourier transform

$$\hat{P}(k,t) = \int\limits_{-\infty}^{\infty} dx\ P(x,t)\,e^{-ikx}\ . \tag{2.78}$$

Expressed in terms of independent variables k and t, one finds that the FPE becomes

$$\partial_t \hat{P} + \beta k\,\partial_k \hat{P} = -Dk^2\hat{P}\ . \tag{2.79}$$

This is known as a *quasilinear partial differential equation*, and a general method of solution for such equations is the *method of characteristics*, which is briefly reviewed in appendix §2.10. A quasilinear PDE in N independent variables can be transformed into $N+1$ coupled ODEs. Applying the method to Eqn. 2.79, one finds

$$\hat{P}(k,t) = \hat{P}\big(k\,e^{-\beta t}, t=0\big)\ \exp\left\{-\frac{D}{2\beta}\left(1-e^{-2\beta t}\right)k^2\right\}\ . \tag{2.80}$$

Suppose $P(x,0) = \delta(x-x_0)$, in which case $\hat{P}(k,0) = e^{-ikx_0}$. We may now apply the inverse Fourier transform to obtain

$$P(x,t) = \sqrt{\frac{\beta}{2\pi D}\cdot\frac{1}{1-e^{-2\beta t}}}\ \exp\left\{-\frac{\beta}{2D}\,\frac{\left(x-x_0\,e^{-\beta t}\right)^2}{1-e^{-2\beta t}}\right\}\ . \tag{2.81}$$

Taking the limit $t\to\infty$, we obtain the asymptotic distribution

$$P(x,t\to\infty) = \sqrt{\frac{\beta}{2\pi D}}\,e^{-\beta x^2/2D}\ , \tag{2.82}$$

which is a Gaussian centered at $x=0$, with standard deviation $\sigma = \sqrt{D/\beta}$.

Physically, the drift term $F_1(x) = -\beta x$ arises when the particle is confined to a harmonic well. The equation of motion is then $\ddot{x} + \gamma\dot{x} + \omega_0^2 x = \eta$, which is discussed in the appendix, §2.8. If we average over the random forcing, then setting the acceleration to zero yields the local drift velocity $v_{\rm drift} = -\omega_0^2\,x/\gamma$, hence $\beta = \omega_0^2/\gamma$. Solving by Laplace transform, one has $L(z) = z^2 + \gamma z + \omega_0^2$, with roots $z_\pm = -\frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} - \omega_0^2}$, and

$$K(s) = \frac{e^{z_+ s} - e^{z_- s}}{z_+ - z_-}\,\Theta(s)\ . \tag{2.83}$$

Note that $\mathrm{Re}\,(z_\pm) < 0$. Plugging this result into Eqn. 2.61 and integrating, we find

$$\lim_{t\to\infty} M(t,t) = \frac{\gamma\Gamma}{\omega_0^2}\ , \tag{2.84}$$

hence the asymptotic distribution is

$$P(x, t\to\infty) = \sqrt{\frac{\gamma\omega_0^2}{2\pi\Gamma}}\, e^{-\gamma\omega_0^2 x^2/2\Gamma}\ . \tag{2.85}$$

Comparing with Eqn. 2.82, we once again find $D = \Gamma/2\gamma^2$. Does the Langevin particle in a harmonic well describe an Ornstein-Uhlenbeck process for finite t? It does in the limit $\gamma\to\infty$, $\omega_0\to\infty$, $\Gamma\to\infty$, with $\beta = \omega_0^2/\gamma$ and $D = \Gamma/2\gamma^2$ finite. In this limit, one has $M(t,t) = \beta^{-1}D\left(1 - e^{-\beta t}\right)$. For $\gamma < \infty$, the velocity relaxation time is finite, and on time scales shorter than γ^{-1} the path $x(t)$ is not Markovian.

In the Ornstein-Uhlenbeck model, drift would like to collapse the distribution to a delta-function at $x = 0$, whereas diffusion would like to spread the distribution infinitely thinly over all space. In that sense, both terms represent extremist inclinations. Yet in the limit $t\to\infty$, drift and diffusion gracefully arrive at a grand compromise, with neither achieving its ultimate goal. The asymptotic distribution is centered about $x = 0$, but has a finite width. There is a lesson here for the United States Congress, if only they understood math.

2.5 The Master Equation

Let $P_i(t)$ be the probability that the system is in a quantum or classical state i at time t. Then write

$$\frac{dP_i}{dt} = \sum_j \left(W_{ij}\,P_j - W_{ji}\,P_i\right)\ , \tag{2.86}$$

where W_{ij} is the rate at which j makes a transition to i. This is known as the *Master equation*. Note that we can recast the Master equation in the form

$$\frac{dP_i}{dt} = -\sum_j \Gamma_{ij}\,P_j\ , \tag{2.87}$$

with

$$\Gamma_{ij} = \begin{cases} -W_{ij} & \text{if } i\neq j \\ \sum_k{}' W_{kj} & \text{if } i = j\ , \end{cases} \tag{2.88}$$

where the prime on the sum indicates that $k = j$ is to be excluded. The constraints on the W_{ij} are that $W_{ij}\geq 0$ for all i, j, and we may take $W_{ii}\equiv 0$ (no sum on i). Fermi's Golden Rule of quantum mechanics says that

$$W_{ij} = \frac{2\pi}{\hbar}\left|\langle\, i\,|\,\hat V\,|\,j\,\rangle\right|^2 \rho(E_j)\ , \tag{2.89}$$

where $\hat H_0\,\big|\,i\,\big\rangle = E_i\,\big|\,i\,\big\rangle$, $\hat V$ is an additional potential which leads to transitions, and $\rho(E_i)$ is the density of final states at energy E_i. The fact that $W_{ij}\geq 0$ means that if each $P_i(t=0)\geq 0$, then $P_i(t)\geq 0$ for all $t\geq 0$. To see this, suppose that at some time $t > 0$ one of the probabilities P_i is crossing zero and about to become negative. But then eqn. 2.86 says that $\dot P_i(t) = \sum_j W_{ij}P_j(t)\geq 0$. So $P_i(t)$ can never become negative.

2.5.1 Equilibrium distribution and detailed balance

If the transition rates W_{ij} are themselves time-independent, then we may formally write

$$P_i(t) = \left(e^{-\Gamma t}\right)_{ij} P_j(0) \ . \tag{2.90}$$

Here we have used the Einstein 'summation convention' in which repeated indices are summed over (in this case, the j index). Note that

$$\sum_i \Gamma_{ij} = 0 \ , \tag{2.91}$$

which says that the total probability $\sum_i P_i$ is conserved:

$$\frac{d}{dt}\sum_i P_i = -\sum_{i,j} \Gamma_{ij}\, P_j = -\sum_j \left(P_j \sum_i \Gamma_{ij} \right) = 0 \ . \tag{2.92}$$

We conclude that $\vec{\phi} = (1, 1, \ldots, 1)$ is a left eigenvector of Γ with eigenvalue $\lambda = 0$. The corresponding right eigenvector, which we write as $P_i^{\rm eq}$, satisfies $\Gamma_{ij} P_j^{\rm eq} = 0$, and is a stationary (*i.e.* time independent) solution to the Master equation. Generally, there is only one right/left eigenvector pair corresponding to $\lambda = 0$, in which case any initial probability distribution $P_i(0)$ converges to $P_i^{\rm eq}$ as $t \to \infty$.

In equilibrium, the net rate of transitions into a state $|\, i \,\rangle$ is equal to the rate of transitions out of $|\, i \,\rangle$. If, for each state $|\, j \,\rangle$ the transition rate from $|\, i \,\rangle$ to $|\, j \,\rangle$ is equal to the transition rate from $|\, j \,\rangle$ to $|\, i \,\rangle$, we say that the rates satisfy the condition of *detailed balance*. In other words,

$$W_{ij}\, P_j^{\rm eq} = W_{ji}\, P_i^{\rm eq}. \tag{2.93}$$

Assuming $W_{ij} \neq 0$ and $P_j^{\rm eq} \neq 0$, we can divide to obtain

$$\frac{W_{ji}}{W_{ij}} = \frac{P_j^{\rm eq}}{P_i^{\rm eq}} \ . \tag{2.94}$$

Note that detailed balance is a stronger condition than that required for a stationary solution to the Master equation.

If $\Gamma = \Gamma^{\rm t}$ is symmetric, then the right eigenvectors and left eigenvectors are transposes of each other, hence $P^{\rm eq} = 1/N$, where N is the dimension of Γ. The system then satisfies the conditions of detailed balance. See Appendix II (§2.5.3) for an example of this formalism applied to a model of radioactive decay.

2.5.2 Boltzmann's H-theorem

Suppose for the moment that Γ is a symmetric matrix, *i.e.* $\Gamma_{ij} = \Gamma_{ji}$. Then construct the function

$$H(t) = \sum_i P_i(t)\, \ln P_i(t) \ . \tag{2.95}$$

Then

$$\begin{aligned} \frac{dH}{dt} &= \sum_i \frac{dP_i}{dt}\left(1 + \ln P_i\right) = \sum_i \frac{dP_i}{dt}\, \ln P_i \\ &= -\sum_{i,j} \Gamma_{ij}\, P_j \ln P_i \\ &= \sum_{i,j} \Gamma_{ij}\, P_j \big(\ln P_j - \ln P_i \big) \ , \end{aligned} \tag{2.96}$$

where we have used $\sum_i \Gamma_{ij} = 0$. Now switch $i \leftrightarrow j$ in the above sum and add the terms to get

$$\frac{dH}{dt} = \frac{1}{2}\sum_{i,j} \Gamma_{ij}\big(P_i - P_j\big)\left(\ln P_i - \ln P_j\right) . \tag{2.97}$$

Note that the $i = j$ term does not contribute to the sum. For $i \neq j$ we have $\Gamma_{ij} = -W_{ij} \leq 0$, and using the result

$$(x - y)\,(\ln x - \ln y) \geq 0 , \tag{2.98}$$

we conclude

$$\frac{dH}{dt} \leq 0 . \tag{2.99}$$

In equilibrium, $P_i^{\rm eq}$ is a constant, independent of i. We write

$$P_i^{\rm eq} = \frac{1}{\Omega} \quad , \quad \Omega = \sum_i 1 \quad \Longrightarrow \quad H = -\ln\Omega . \tag{2.100}$$

If $\Gamma_{ij} \neq \Gamma_{ji}$, we can still prove a version of the H-theorem. Define a new symmetric matrix

$$\overline{W}_{ij} \equiv W_{ij}\,P_j^{\rm eq} = W_{ji}\,P_i^{\rm eq} = \overline{W}_{ji} , \tag{2.101}$$

and the generalized H-function,

$$H(t) \equiv \sum_i P_i(t)\,\ln\!\left(\frac{P_i(t)}{P_i^{\rm eq}}\right) . \tag{2.102}$$

Then

$$\frac{dH}{dt} = -\frac{1}{2}\sum_{i,j} \overline{W}_{ij}\left(\frac{P_i}{P_i^{\rm eq}} - \frac{P_j}{P_j^{\rm eq}}\right)\left[\ln\!\left(\frac{P_i}{P_i^{\rm eq}}\right) - \ln\!\left(\frac{P_j}{P_j^{\rm eq}}\right)\right] \leq 0 . \tag{2.103}$$

2.5.3 Formal solution to the Master equation

Recall the Master equation $\dot{P}_i = -\Gamma_{ij}\,P_j$. The matrix Γ_{ij} is real but not necessarily symmetric. For such a matrix, the left eigenvectors ϕ_i^α and the right eigenvectors ψ_j^β are not the same: general different:

$$\begin{aligned} \phi_i^\alpha\,\Gamma_{ij} &= \lambda_\alpha\,\phi_j^\alpha \\ \Gamma_{ij}\,\psi_j^\beta &= \lambda_\beta\,\psi_i^\beta . \end{aligned} \tag{2.104}$$

Note that the eigenvalue equation for the right eigenvectors is $\Gamma\psi = \lambda\psi$ while that for the left eigenvectors is $\Gamma^{\rm t}\phi = \lambda\phi$. The characteristic polynomial is the same in both cases:

$$F(\lambda) \equiv \det\left(\lambda - \Gamma\right) = \det\left(\lambda - \Gamma^{\rm t}\right) , \tag{2.105}$$

which means that the left and right eigenvalues are the same. Note also that $\big[F(\lambda)\big]^* = F(\lambda^*)$, hence the eigenvalues are either real or appear in complex conjugate pairs. Multiplying the eigenvector equation for ϕ^α on the right by ψ_j^β and summing over j, and multiplying the eigenvector equation for ψ^β on the left by ϕ_i^α and summing over i, and subtracting the two results yields

$$\big(\lambda_\alpha - \lambda_\beta\big)\,\big\langle\,\phi^\alpha\,\big|\,\psi^\beta\,\big\rangle = 0 , \tag{2.106}$$

where the inner product is

$$\big\langle\,\phi\,\big|\,\psi\,\big\rangle = \sum_i \phi_i\,\psi_i . \tag{2.107}$$

We can now demand

$$\left\langle \, \phi^\alpha \, \middle| \, \psi^\beta \, \right\rangle = \delta_{\alpha\beta} \ , \tag{2.108}$$

in which case we can write

$$\Gamma = \sum_\alpha \lambda_\alpha \left| \, \psi^\alpha \, \right\rangle \! \left\langle \, \phi^\alpha \, \right| \qquad \Longleftrightarrow \qquad \Gamma_{ij} = \sum_\alpha \lambda_\alpha \, \psi_i^\alpha \, \phi_j^\alpha \ . \tag{2.109}$$

We have seen that $\vec{\phi} = (1, 1, \ldots, 1)$ is a left eigenvector with eigenvalue $\lambda = 0$, since $\sum_i \Gamma_{ij} = 0$. We do not know *a priori* the corresponding right eigenvector, which depends on other details of Γ_{ij}. Now let's expand $P_i(t)$ in the right eigenvectors of Γ, writing

$$P_i(t) = \sum_\alpha C_\alpha(t) \, \psi_i^\alpha \ . \tag{2.110}$$

Then

$$\begin{aligned} \frac{dP_i}{dt} &= \sum_\alpha \frac{dC_\alpha}{dt} \, \psi_i^\alpha \\ &= -\Gamma_{ij} \, P_j = -\sum_\alpha C_\alpha \, \Gamma_{ij} \, \psi_j^\alpha = -\sum_\alpha \lambda_\alpha \, C_\alpha \, \psi_i^\alpha \ , \end{aligned} \tag{2.111}$$

and linear independence of the eigenvectors $\left| \, \psi^\alpha \, \right\rangle$ allows us to conclude

$$\frac{dC_\alpha}{dt} = -\lambda_\alpha \, C_\alpha \qquad \Longrightarrow \qquad C_\alpha(t) = C_\alpha(0) \, e^{-\lambda_\alpha t} \ . \tag{2.112}$$

Hence, we can write

$$P_i(t) = \sum_\alpha C_\alpha(0) \, e^{-\lambda_\alpha t} \, \psi_i^\alpha \ . \tag{2.113}$$

It is now easy to see that $\mathrm{Re}\,(\lambda_\alpha) \geq 0$ for all λ, or else the probabilities will become negative. For suppose $\mathrm{Re}\,(\lambda_\alpha) < 0$ for some α. Then as $t \to \infty$, the sum in eqn. 2.113 will be dominated by the term for which λ_α has the largest negative real part; all other contributions will be subleading. But we must have $\sum_i \psi_i^\alpha = 0$ since $\left| \, \psi^\alpha \, \right\rangle$ must be orthogonal to the left eigenvector $\vec{\phi}^{\alpha=0} = (1, 1, \ldots, 1)$. Therefore, at least one component of ψ_i^α (*i.e.* for some value of i) must have a negative real part, which means a negative probability![8] As we have already proven that an initial nonnegative distribution $\{P_i(t = 0)\}$ will remain nonnegative under the evolution of the Master equation, we conclude that $P_i(t) \to P_i^{\rm eq}$ as $t \to \infty$, relaxing to the $\lambda = 0$ right eigenvector, with $\mathrm{Re}\,(\lambda_\alpha) \geq 0 \; \forall \; \alpha$.

Poisson process

Consider the *Poisson process*, for which

$$W_{mn} = \begin{cases} \lambda & \text{if } m = n + 1 \\ 0 & \text{if } m \neq n + 1 \ . \end{cases} \tag{2.114}$$

We then have

$$\frac{dP_n}{dt} = \lambda \big(P_{n-1} - P_n \big) \ . \tag{2.115}$$

[8] Since the probability $P_i(t)$ is real, if the eigenvalue with the smallest (*i.e.* largest negative) real part is complex, there will be a corresponding complex conjugate eigenvalue, and summing over all eigenvectors will result in a real value for $P_i(t)$.

The *generating function* $P(z,t) = \sum_{n=0}^\infty z^n P_n(t)$ then satisfies

$$\frac{\partial P}{\partial t} = \lambda(z-1)\,P \quad \Rightarrow \quad P(z,t) = e^{(z-1)\lambda t}\,P(z,0)\;. \tag{2.116}$$

If the initial distribution is $P_n(0) = \delta_{n,0}$, then

$$P_n(t) = \frac{(\lambda t)^n}{n!}\,e^{-\lambda t}\;, \tag{2.117}$$

which is known as the *Poisson distribution*. If we define $\alpha \equiv \lambda t$, then from $P_n = \alpha^n\,e^{-\alpha}/n!$ we have

$$\langle n^k\rangle = e^{-\alpha}\left(\alpha\frac{\partial}{\partial\alpha}\right)^k e^\alpha\;. \tag{2.118}$$

Thus, $\langle n\rangle = \alpha$, $\langle n^2\rangle = \alpha^2 + \alpha$, *etc.*

Radioactive decay

Consider a group of atoms, some of which are in an excited state which can undergo nuclear decay. Let $P_n(t)$ be the probability that n atoms are excited at some time t. We then model the decay dynamics by

$$W_{mn} = \begin{cases} 0 & \text{if } m \geq n \\ n\gamma & \text{if } m = n-1 \\ 0 & \text{if } m < n-1\;. \end{cases} \tag{2.119}$$

Here, γ is the decay rate of an individual atom, which can be determined from quantum mechanics. The Master equation then tells us

$$\frac{dP_n}{dt} = (n+1)\,\gamma\,P_{n+1} - n\,\gamma\,P_n\;. \tag{2.120}$$

The interpretation here is as follows: let $|\,n\,\rangle$ denote a state in which n atoms are excited. Then $P_n(t) = \big|\langle\,\psi(t)\,|\,n\,\rangle\big|^2$. Then $P_n(t)$ will increase due to spontaneous transitions from $|\,n+1\,\rangle$ to $|\,n\,\rangle$, and will decrease due to spontaneous transitions from $|\,n\,\rangle$ to $|\,n-1\,\rangle$.

The average number of particles in the system is $N(t) = \sum_{n=0}^\infty n\,P_n(t)$. Note that

$$\frac{dN}{dt} = \sum_{n=0}^\infty n\Big[(n+1)\,\gamma\,P_{n+1} - n\,\gamma\,P_n\Big] = -\gamma\sum_{n=0}^\infty n\,P_n = -\gamma\,N\;. \tag{2.121}$$

Thus, $N(t) = N(0)\,e^{-\gamma t}$. The relaxation time is $\tau = \gamma^{-1}$, and the equilibrium distribution is $P_n^{\rm eq} = \delta_{n,0}$, which satisfies detailed balance.

Making use again of the generating function $P(z,t) = \sum_{n=0}^\infty z^n\,P_n(t)$, we derive the PDE

$$\frac{\partial P}{\partial t} = \gamma\sum_{n=0}^\infty z^n\Big[(n+1)\,P_{n+1} - n\,P_n\Big] = \gamma\,\frac{\partial P}{\partial z} - \gamma z\,\frac{\partial P}{\partial z}\;. \tag{2.122}$$

Thus, we have $\partial_t P = \gamma(1-z)\,\partial_z P$, which is solved by any function $f(\xi)$, where $\xi = \gamma t - \ln(1-z)$. Thus, we can write $P(z,t) = f\big(\gamma t - \ln(1-z)\big)$. Setting $t=0$ we have $P(z,0) = f\big(-\ln(1-z)\big)$, whence $f(u) = P(1-e^{-u},0)$ is now given in terms of the initial distribution $P(z,t=0)$. Thus, the full solution for $P(z,t)$ is

$$P(z,t) = P\big(1 + (z-1)\,e^{-\gamma t},\,0\big)\;. \tag{2.123}$$

The total probability is $P(z{=}1,t) = \sum_{n=0}^\infty P_n$, which clearly is conserved: $P(1,t) = P(1,0)$. The average particle number is then $N(t) = \partial_z\,P(z,t)\big|_{z=1} = e^{-\gamma t}\,P(1,0) = e^{-\gamma t}\,N(0)$.

2.6 Formal Theory of Stochastic Processes

Here we follow the presentation in chapter 3 in the book by C. Gardiner. Given a time-dependent random variable $\boldsymbol{X}(t)$, we define the probability distribution

$$P(\boldsymbol{x},t) = \big\langle \delta\big(\boldsymbol{x} - \boldsymbol{X}(t)\big)\big\rangle \ , \tag{2.124}$$

where the average is over different realizations of the random process. $P(\boldsymbol{x},t)$ is a density with units L^{-d}. This distribution is normalized according to $\int d\boldsymbol{x}\, P(\boldsymbol{x},t) = 1$, where $d\boldsymbol{x} = d^d x$ is the differential for the spatial volume, and does not involve time. If we integrate over some region A, we obtain

$$P_A(t) = \int\limits_A d\boldsymbol{x}\, P(\boldsymbol{x},t) = \text{probability that } \boldsymbol{X}(t) \in A \ . \tag{2.125}$$

We define the *joint probability distributions* as follows:

$$P(\boldsymbol{x}_1,t_1\,;\,\boldsymbol{x}_2,t_2\,;\,\ldots\,;\,\boldsymbol{x}_N,t_N) = \Big\langle \delta\big(\boldsymbol{x}_1 - \boldsymbol{X}(t_1)\big)\cdots\delta\big(\boldsymbol{x}_N - \boldsymbol{X}(t_N)\big)\Big\rangle \ . \tag{2.126}$$

From the joint probabilities we may form *conditional probability distributions*

$$P(\boldsymbol{x}_1,t_1\,;\,\boldsymbol{x}_2,t_2\,;\,\ldots\,;\,\boldsymbol{x}_N,t_N\,|\,\boldsymbol{y}_1,\tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M,\tau_M) = \frac{P(\boldsymbol{x}_1,t_1\,;\,\ldots\,;\,\boldsymbol{x}_N,t_N\,;\,\boldsymbol{y}_1,\tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M,\tau_M)}{P(\boldsymbol{y}_1,\tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M,\tau_M)} \ . \tag{2.127}$$

Although the times can be in any order, by convention we order them so they decrease from left to right:

$$t_1 > \cdots > t_N > \tau_1 > \cdots \tau_M \ . \tag{2.128}$$

2.6.1 Markov processes

In a *Markov process*, any conditional probability is determined by its *most recent condition*. Thus,

$$P(\boldsymbol{x}_1,t_1\,;\,\boldsymbol{x}_2,t_2\,;\,\ldots\,;\,\boldsymbol{x}_N,t_N\,|\,\boldsymbol{y}_1,\tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M,\tau_M) = P(\boldsymbol{x}_1,t_1\,;\,\boldsymbol{x}_2,t_2\,;\,\ldots\,;\,\boldsymbol{x}_N,t_N\,|\,\boldsymbol{y}_1,\tau_1) \ , \tag{2.129}$$

where the ordering of the times is as in Eqn. 2.128. This definition entails that *all* probabilities may be constructed from $P(\boldsymbol{x},t)$ and from the conditional $P(\boldsymbol{x},t\,|\,\boldsymbol{y},\tau)$. Clearly $P(\boldsymbol{x}_1,t_1\,;\,\boldsymbol{x}_2,t_2) = P(\boldsymbol{x}_1,t_1\,|\,\boldsymbol{x}_2,t_2)\,P(\boldsymbol{x}_2,t_2)$. At the next level, we have

$$\begin{aligned} P(\boldsymbol{x}_1,t_1\,;\,\boldsymbol{x}_2,t_2\,;\,\boldsymbol{x}_3,t_3) &= P(\boldsymbol{x}_1,t_1\,|\,\boldsymbol{x}_2,t_2\,;\,\boldsymbol{x}_3,t_3)\,P(\boldsymbol{x}_2,t_2\,;\,\boldsymbol{x}_3,t_3) \\ &= P(\boldsymbol{x}_1,t_1\,|\,\boldsymbol{x}_2,t_2)\,P(\boldsymbol{x}_2,t_2\,|\,\boldsymbol{x}_3,t_3)\,P(\boldsymbol{x}_3,t_3) \ . \end{aligned}$$

Proceeding thusly, we have

$$P(\boldsymbol{x}_1,t_1\,;\,\ldots\,;\,\boldsymbol{x}_N,t_N) = P(\boldsymbol{x}_1,t_1\,|\,\boldsymbol{x}_2,t_2)\,P(\boldsymbol{x}_2,t_2\,|\,\boldsymbol{x}_3,t_3)\cdots P(\boldsymbol{x}_{N-1},t_{N-1}\,|\,\boldsymbol{x}_N,t_N)\,P(\boldsymbol{x}_N,t_N) \ , \tag{2.130}$$

so long as $t_1 > t_2 > \ldots > t_N$.

Chapman-Kolmogorov equation

The probability density $P(\boldsymbol{x}_1,t_1)$ can be obtained from the joint probability density $P(\boldsymbol{x}_1,t_1\,;\,\boldsymbol{x}_2,t_2)$ by integrating over $\boldsymbol{x}_2$:

$$P(\boldsymbol{x}_1,t_1) = \int d\boldsymbol{x}_2\, P(\boldsymbol{x}_1,t_1\,;\,\boldsymbol{x}_2,t_2) = \int d\boldsymbol{x}_2\, P(\boldsymbol{x}_1,t_1\,|\,\boldsymbol{x}_2,t_2)\,P(\boldsymbol{x}_2,t_2) \ . \tag{2.131}$$

Similarly[9],

$$P(\boldsymbol{x}_1, t_1 \,|\, \boldsymbol{x}_3, t_3) = \int d\boldsymbol{x}_2 \, P(\boldsymbol{x}_1, t_1 \,|\, \boldsymbol{x}_2, t_2 \,;\, \boldsymbol{x}_3, t_3) \, P(\boldsymbol{x}_2, t_2 \,|\, \boldsymbol{x}_3, t_3) \ . \tag{2.132}$$

For Markov processes, then,

$$P(\boldsymbol{x}_1, t_1 \,|\, \boldsymbol{x}_3, t_3) = \int d\boldsymbol{x}_2 \, P(\boldsymbol{x}_1, t_1 \,|\, \boldsymbol{x}_2, t_2) \, P(\boldsymbol{x}_2, t_2 \,|\, \boldsymbol{x}_3, t_3) \ . \tag{2.133}$$

For discrete spaces, we have $\int d\boldsymbol{x} \to \sum_{\boldsymbol{x}}$, and $\sum_{\boldsymbol{x}_2} P(\boldsymbol{x}_1, t_1 \,|\, \boldsymbol{x}_2, t_2) \, P(\boldsymbol{x}_2, t_2 \,|\, \boldsymbol{x}_3, t_3)$ is a matrix multiplication.

Do Markov processes exist in nature and are they continuous?

A random walk in which each step is independently and identically distributed is a Markov process. Consider now the following arrangement. You are given a bag of marbles, an initial fraction p_0 of which are red, q_0 of which are green, and r_0 of which are blue, with $p_0 + q_0 + r_0 = 1$. Let $\sigma_j = +1$, 0, or -1 according to whether the $j^{\rm th}$ marble selected is red, green, or blue, respectively, and define $X_n = \sum_{j=1}^n \sigma_j$, which would correspond to the position of a random walker who steps either to the right ($\sigma_j = +1$), remain stationary ($\sigma_j = 0$), or steps left ($\sigma_j = -1$) during each discrete time interval. If the bag is infinite, then $\{X_1, X_2, \ldots\}$ is a Markov process. The probability for $\sigma_j = +1$ remains at $p = p_0$ and is unaffected by the withdrawal of any finite number of marbles from the bag. But if the contents of the bag are *finite*, then the probability p changes with discrete time, and in such a way that cannot be determined from the instantaneous value of X_n alone. Note that if there were only two colors of marbles, and $\sigma_j \in \{+1\,,\,-1\}$, then given $X_0 = 0$ and knowledge of the initial number of marbles in the bag, specifying X_n tells us everything we need to know about the composition of the bag at time n. But with three possibilities $\sigma_j \in \{+1\,,\,0\,,\,-1\}$ we need to know the entire history in order to determine the current values of p, q, and r. The reason is that the sequences 0000, $001\overline{1}$, $\overline{1}11\overline{1}$ (with $\overline{1} \equiv -1$) all have the same effect on the displacement X, but result in a different composition of marbles remaining in the bag.

In physical systems, processes we might model as random have a finite correlation time. We saw above that the correlator of the random force $\eta(t)$ in the Langevin equation is written $\big\langle \eta(t)\,\eta(t+s) \big\rangle = \phi(s)$, where $\phi(s)$ decays to zero on a time scale τ_ϕ. For time differences $|s| < \tau_\phi$, the system is not Markovian. In addition, the system itself may exhibit some memory. For example, in the Langevin equation $\dot u + \gamma u = \eta(t)$, there is a time scale γ^{-1} over which the variable $p(t)$ forgets its previous history. Still, if $\tau_\phi = 0$, $u(t)$ is a Markov process, because the equation is first order and therefore only the most recent condition is necessary in order to integrate forward from some past time $t = t_0$ to construct the statistical ensemble of functions $u(t)$ for $t > t_0$. For second order equations, such as $\ddot x + \gamma \dot x = \eta(t)$, two initial conditions are required, hence diffusion paths $X(t)$ are only Markovian on time scales beyond γ^{-1}, over which the memory of the initial velocity is lost. More generally, if $\boldsymbol{\varphi}$ is an N-component vector in phase space, and

$$\frac{d\varphi_i}{dt} = A_i(\boldsymbol{\varphi}, t) + B_{ij}(\boldsymbol{\varphi}, t)\,\eta_j(t) \ , \tag{2.134}$$

where we may choose $\big\langle \eta_i(t)\,\eta_j(t') \big\rangle = \delta_{ij}\,\delta(t-t')$, then the path $\boldsymbol{\varphi}(t)$ is a Markov process.

While a random variable $\boldsymbol{X}(t)$ may take values in a continuum, as a function of time it may still exhibit discontinuous jumps. That is to say, even though time t may evolve continuously, the sample paths $\boldsymbol{X}(t)$ may be discontinuous. As an example, consider the Brownian motion of a particle moving in a gas or fluid. On the scale of the autocorrelation time, the velocity changes discontinuously, while the position $\boldsymbol{X}(t)$ evolves continuously (although not smoothly). The condition that sample paths $\boldsymbol{X}(t)$ evolve continuously is known as the *Lindeberg condition*,

$$\lim_{\tau \to 0} \frac{1}{\tau} \int\limits_{|\boldsymbol{x}-\boldsymbol{y}|>\varepsilon} d\boldsymbol{y} \, P(\boldsymbol{y}, t+\tau \,|\, \boldsymbol{x}, t) = 0 \ . \tag{2.135}$$

[9]Because $P(\boldsymbol{x}_1, t_1 \,;\, \boldsymbol{x}_2, t_3 \,|\, \boldsymbol{x}_3, t_3) = \big[P(\boldsymbol{x}_1, t_1 \,;\, \boldsymbol{x}_2, t_2 \,;\, \boldsymbol{x}_3, t_3)/P(\boldsymbol{x}_2, t_2 \,;\, \boldsymbol{x}_3, t_3)\big] \cdot \big[P(\boldsymbol{x}_2, t_2 \,;\, \boldsymbol{x}_3, t_3)/P(\boldsymbol{x}_3, t_3)\big]$.

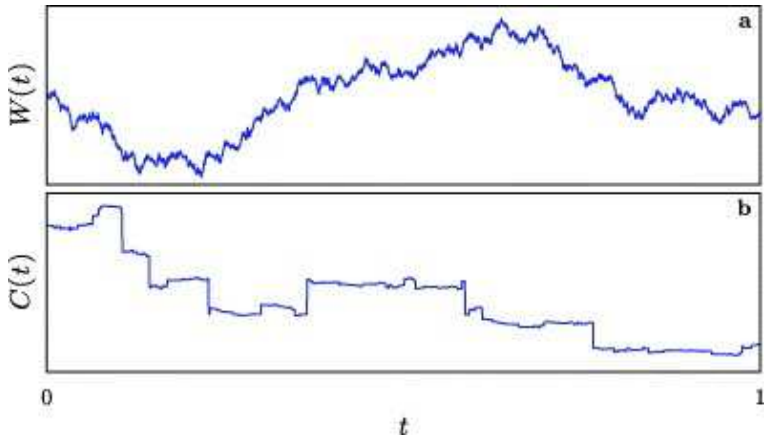

Figure 2.3: (a) Wiener process sample path $W(t)$. (b) Cauchy process sample path $C(t)$. From K. Jacobs and D. A. Steck, *New J. Phys.* **13**, 013016 (2011).

If this condition is satisfied, then the sample paths $\boldsymbol{X}(t)$ are continuous with probability one. Two examples:

(1) **Wiener process**: As we shall discuss below, this is a pure diffusion process with no drift or jumps, with

$$P(x,t\,|\,x',t') = \frac{1}{\sqrt{4\pi D|t-t'|}}\,\exp\left(-\frac{(x-x')^2}{4D|t-t'|}\right) \tag{2.136}$$

in one space dimension. The Lindeberg condition is satisfied, and the sample paths $\boldsymbol{X}(t)$ are continuous.

(2) **Cauchy process**: This is a process in which sample paths exhibit finite jumps, and hence are not continuous. In one space dimension,

$$P(x,t\,|\,x',t') = \frac{|t-t'|}{\pi\big[(x-x')^2+(t-t')^2\big]}\ . \tag{2.137}$$

Note that in both this case and the Wiener process described above, we have $\lim_{t-t'\to 0} P(xt\,|\,x't') = \delta(x-x')$. However in this example the Lindeberg condition is not satisfied.

To simulate, given $\boldsymbol{x}_n = \boldsymbol{X}(t=n\tau)$, choose $\boldsymbol{y}\in\mathcal{D}_b(\boldsymbol{x}_n)$, where $\mathcal{D}_b(\boldsymbol{x}_n)$ is a ball of radius $b>\varepsilon$ centered at $\boldsymbol{x}_n$. Then evaluate the probability $p\equiv P(\boldsymbol{y},(n+1)\tau\,|\,\boldsymbol{x},n\tau)$. If p exceeds a random number drawn from a uniform distribution on $[0,1]$, accept and set $\boldsymbol{x}_{n+1} = \boldsymbol{X}\big((n+1)\tau\big) = \boldsymbol{y}$. Else reject and choose a new $\boldsymbol{y}$ and proceed as before.

2.6.2 Martingales

A *Martingale* is a stochastic process for which the *conditional average* of the random variable $\boldsymbol{X}(t)$ does not change from its most recent condition. That is,

$$\big\langle \boldsymbol{x}(t)\,\big|\,\{\boldsymbol{y}_1\,\tau_1\,;\,\boldsymbol{y}_2,\tau_2\,;\,\ldots\,;\,\boldsymbol{y}_M,\tau_M\}\big\rangle = \int d\boldsymbol{x}\,P(\boldsymbol{x},t\,|\,\boldsymbol{y}_1,\tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M,\tau_M)\,\boldsymbol{x} = \boldsymbol{y}_1\ . \tag{2.138}$$

In this sense, a Martingale is a stochastic process which represents a 'fair game'. Not every Martingale is a Markov process, and not every Markov process is a Martingale. The Wiener process is a Martingale.

One very important fact about Martingales, which we will here derive in $d = 1$ dimension. For $t_1 > t_2$,

$$\begin{aligned}\big\langle x(t_1)\,x(t_2)\big\rangle &= \int\!dx_1\!\int\!dx_2\;P(x_1,t_1\,;\,x_2,t_2)\,x\,x_2 = \int\!dx_1\!\int\!dx_2\;P(x_1,t_1\,;\,x_2,t_2)\,P(x_2,t_2)\,x_1\,x_2 \\ &= \int\!dx_2\;P(x_2,t_2)\,x_2\!\int\!dx_1\;P(x_1,t_1\,|\,x_2,t_2)\,x_1 = \int\!dx_2\;P(x_2,t_2)\,x_2^2 \\ &= \big\langle x^2(t_2)\big\rangle\;.\end{aligned} \tag{2.139}$$

One can further show that, for $t_2 > t_2 > t_3$,

$$\Big\langle \big[x(t_1)-x(t_2)\big]\big[x(t_2)-x(t_3)\big]\Big\rangle = 0\;, \tag{2.140}$$

which says that *at the level of pair correlations*, past performance provides no prediction of future results.

2.6.3 Differential Chapman-Kolmogorov equations

Suppose the following conditions apply:

$$|\boldsymbol{y}-\boldsymbol{x}|>\varepsilon \quad\Longrightarrow\quad \lim_{\tau\to 0}\,\frac{1}{\tau}\,P(\boldsymbol{y},t+\tau\,|\,\boldsymbol{x},t) = W(\boldsymbol{y}\,|\,\boldsymbol{x},t) \tag{2.141}$$

$$\lim_{\tau\to 0}\,\frac{1}{\tau}\int\limits_{|\boldsymbol{y}-\boldsymbol{x}|<\varepsilon}\!\!d\boldsymbol{y}\;(y_\mu - x_\mu)\,P(\boldsymbol{y},t+\tau\,|\,\boldsymbol{x},t) = A_\mu(\boldsymbol{x},t)+\mathcal{O}(\varepsilon) \tag{2.142}$$

$$\lim_{\tau\to 0}\,\frac{1}{\tau}\int\limits_{|\boldsymbol{y}-\boldsymbol{x}|<\varepsilon}\!\!d\boldsymbol{y}\;(y_\mu - x_\mu)\,(y_\nu - x_\nu)\,P(\boldsymbol{y},t+\tau\,|\,\boldsymbol{x},t) = B_{\mu\nu}(\boldsymbol{x},t)+\mathcal{O}(\varepsilon)\;, \tag{2.143}$$

where the last two conditions hold uniformly in $\boldsymbol{x}$, t, and ε. Then following §3.4.1 and §3.6 of Gardiner, one obtains the *forward differential Chapman-Kolmogorov equation* (DCK+),

$$\begin{aligned}\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t} &= -\sum_\mu \frac{\partial}{\partial x_\mu}\Big[A_\mu(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big] + \frac{1}{2}\sum_{\mu,\nu}\frac{\partial^2}{\partial x_\mu\,\partial x_\nu}\Big[B_{\mu\nu}(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big] \\ &\qquad + \int\!d\boldsymbol{y}\,\Big[W(\boldsymbol{x}\,|\,\boldsymbol{y},t)\,P(\boldsymbol{y},t\,|\,\boldsymbol{x}',t') - W(\boldsymbol{y}\,|\,\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big]\;,\end{aligned} \tag{2.144}$$

and the *backward differential Chapman-Kolmogorov equation* (DCK−),

$$\begin{aligned}\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t'} &= -\sum_\mu A_\mu(\boldsymbol{x}',t')\,\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial x'_\mu} + \frac{1}{2}\sum_{\mu,\nu}B_{\mu\nu}(\boldsymbol{x}',t')\,\frac{\partial^2 P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial x'_\mu\,\partial x'_\nu} \\ &\qquad + \int\!d\boldsymbol{y}\,W(\boldsymbol{y}\,|\,\boldsymbol{x}',t')\Big[P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t') - P(\boldsymbol{x},t\,|\,\boldsymbol{y},t')\Big]\;.\end{aligned} \tag{2.145}$$

Note that the Lindeberg condition requires that

$$\lim_{\tau\to 0}\,\frac{1}{\tau}\int\limits_{|\boldsymbol{x}-\boldsymbol{y}|>\varepsilon}\!\!d\boldsymbol{y}\;P(\boldsymbol{y},t+\tau\,|\,\boldsymbol{x},t) = \int\limits_{|\boldsymbol{x}-\boldsymbol{y}|>\varepsilon}\!\!d\boldsymbol{y}\;W(\boldsymbol{y}\,|\,\boldsymbol{x},t) = 0\;, \tag{2.146}$$

which must hold for any $\varepsilon > 0$. Taking the limit $\varepsilon \to 0$, we conclude[10] $W(\boldsymbol{y}\,|\,\boldsymbol{x},t) = 0$ *if* the Lindeberg condition is satisfied. If there are any jump processes, *i.e.* if $W(\boldsymbol{y}\,|\,\boldsymbol{x},t)$ does not identically vanish for all values of its arguments, then Lindeberg is violated, and the paths are discontinuous.

[10] What about the case $\boldsymbol{y} = \boldsymbol{x}$, which occurs for $\varepsilon = 0$, which is never actually reached throughout the limiting procedure? The quantity $W(\boldsymbol{x}\,|\,\boldsymbol{x},t)$ corresponds to the rate at which the system jumps from $\boldsymbol{x}$ to $\boldsymbol{x}$ at time t, which is not a jump process at all. Note that the contribution from $\boldsymbol{y} = \boldsymbol{x}$ cancels from the DCK± equations. In other words, we can set $W(\boldsymbol{x}\,|\,\boldsymbol{x},t) \equiv 0$.

Some applications:

(1) **Master equation**: If $A_\mu(\boldsymbol{x},t)=0$ and $B_{\mu\nu}(\boldsymbol{x},t)=0$, then we have from DCK+,

$$\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t}=\int d\boldsymbol{y}\,\Big[W(\boldsymbol{x}\,|\,\boldsymbol{y},t)\,P(\boldsymbol{y},t\,|\,\boldsymbol{x}',t')-W(\boldsymbol{y}\,|\,\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big]\ . \tag{2.147}$$

Let's integrate this equation over a time interval Δt. Assuming $P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t)=\delta(\boldsymbol{x}-\boldsymbol{x}')$, we have

$$P(\boldsymbol{x},t+\Delta t\,|\,\boldsymbol{x}',t)=\Big[1-\Delta t\int d\boldsymbol{y}\,W(\boldsymbol{y}\,|\,\boldsymbol{x}',t)\Big]\delta(\boldsymbol{x}-\boldsymbol{x}')+W(\boldsymbol{x}\,|\,\boldsymbol{x}',t)\,\Delta t\ . \tag{2.148}$$

Thus,

$$Q(\boldsymbol{x}',t+\Delta t,t)=1-\Delta t\int d\boldsymbol{y}\,W(\boldsymbol{y}\,|\,\boldsymbol{x}',t) \tag{2.149}$$

is the probability for a particle to remain at $\boldsymbol{x}'$ over the interval $\big[t,t+\Delta t\big]$ given that it was at $\boldsymbol{x}'$ at time t. Iterating this relation, we find

$$\begin{aligned}Q(\boldsymbol{x},t,t_0)&=\Big(1-\Lambda(\boldsymbol{x},t-\Delta t)\,\Delta t\Big)\Big(1-\Lambda(\boldsymbol{x},t-2\Delta t)\,\Delta t\Big)\cdots\Big(1-\Lambda(\boldsymbol{x},t_0)\,\Delta t\Big)\,\overbrace{Q(\boldsymbol{x},t_0,t_0)}^{1}\\&=\mathrm{P}\exp\left\{-\int_{t_0}^{t}dt'\,\Lambda(\boldsymbol{x},t')\right\}\ ,\end{aligned} \tag{2.150}$$

where $\Lambda(\boldsymbol{x},t)=\int d\boldsymbol{y}\,W(\boldsymbol{y}\,|\,\boldsymbol{x},t)$ and P is the path ordering operator which places earlier times to the right. The interpretation of the function $W(\boldsymbol{y}\,|\,\boldsymbol{x},t)$ is that it is the probability density *rate* for the random variable $\boldsymbol{X}$ to jump from $\boldsymbol{x}$ to $\boldsymbol{y}$ at time $\boldsymbol{t}$. Thus, the dimensions of $W(\boldsymbol{y}\,|\,\boldsymbol{x},t)$ are $L^{-d}\,T^{-1}$. Such processes are called *jump processes*. For discrete state spaces, the Master equation takes the form

$$\frac{\partial P(\boldsymbol{n},t\,|\,\boldsymbol{n}',t')}{\partial t}=\sum_{\boldsymbol{m}}\Big[W(\boldsymbol{n}\,|\,\boldsymbol{m},t)\,P(\boldsymbol{m},t\,|\,\boldsymbol{n}',t')-W(\boldsymbol{m}\,|\,\boldsymbol{n},t)\,P(\boldsymbol{n},t\,|\,\boldsymbol{n}',t')\Big]\ . \tag{2.151}$$

Here $W(\boldsymbol{n}\,|\,\boldsymbol{m},t)$ has units T^{-1}, and corresponds to the rate of transitions from state $\boldsymbol{m}$ to state $\boldsymbol{n}$ at time t.

(2) **Fokker-Planck equation**: If $W(\boldsymbol{x}\,|\,\boldsymbol{y},t)=0$, DCK+ gives

$$\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t}=-\sum_\mu\frac{\partial}{\partial x_\mu}\Big[A_\mu(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big]+\tfrac{1}{2}\sum_{\mu,\nu}\frac{\partial^2}{\partial x_\mu\,\partial x_\nu}\Big[B_{\mu\nu}(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big]\ , \tag{2.152}$$

which is a more general form of the Fokker-Planck equation we studied in §2.4 above. Defining the average $\big\langle F(\boldsymbol{x},t)\big\rangle=\int d^dx\,F(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')$, via integration by parts we derive

$$\begin{aligned}\frac{d}{dt}\,\big\langle x_\mu\big\rangle&=\big\langle A_\mu\big\rangle\\\frac{d}{dt}\,\big\langle x_\mu\,x_\nu\big\rangle&=\big\langle x_\mu\,A_\nu\big\rangle+\big\langle A_\mu\,x_\nu\big\rangle+\tfrac{1}{2}\,\big\langle D_{\mu\nu}+D_{\nu\mu}\big\rangle\quad.\end{aligned} \tag{2.153}$$

For the case where $A_\mu(\boldsymbol{x},t)$ and $B_{\mu\nu}(\boldsymbol{x},t)$ are constants independent of $\boldsymbol{x}$ and t, we have the solution

$$P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')=\det{}^{-1/2}\big[2\pi B\,\Delta t\big]\,\exp\left\{-\frac{1}{2\,\Delta t}\,\big(\Delta x_\mu-A_\mu\,\Delta t\big)\,B^{-1}_{\mu\nu}\,\big(\Delta x_\nu-A_\nu\,\Delta t\big)\right\}\ , \tag{2.154}$$

where $\Delta\boldsymbol{x} \equiv \boldsymbol{x} - \boldsymbol{x}'$ and $\Delta t \equiv t - t'$. This is normalized so that the integral over $\boldsymbol{x}$ is unity. If we subtract out the *drift* $\boldsymbol{A}\,\Delta t$, then clearly

$$\Big\langle \big(\Delta x_\nu - A_\nu\,\Delta t\big)\,\big(\Delta x_\mu - A_\mu\,\Delta t\big)\Big\rangle = B_{\mu\nu}\,\Delta t\ , \tag{2.155}$$

which is diffusive.

(3) **Liouville equation**: If $W(\boldsymbol{x}\,|\,\boldsymbol{y},t) = 0$ and $B_{\mu\nu}(\boldsymbol{x},t) = 0$, then DCK+ gives

$$\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t} = -\sum_\mu \frac{\partial}{\partial x_\mu}\Big[A_\mu(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big]\ . \tag{2.156}$$

This is Liouville's equation from classical mechanics, also known as the continuity equation. Suppressing the $(\boldsymbol{x}',t')$ variables, the above equation is equivalent to

$$\frac{\partial\varrho}{\partial t} + \boldsymbol{\nabla}\cdot(\varrho\,\boldsymbol{v}) = 0\ , \tag{2.157}$$

where $\varrho(\boldsymbol{x},t) = P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')$ and $\boldsymbol{v}(\boldsymbol{x},t) = \boldsymbol{A}(\boldsymbol{x},t)$. The product of $\boldsymbol{A}$ and P is the current is $\boldsymbol{j} = \varrho\boldsymbol{v}$. To find the general solution, we assume the initial conditions are $P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t) = \delta(\boldsymbol{x}-\boldsymbol{x}')$. Then if $\boldsymbol{x}(t;\boldsymbol{x}')$ is the solution to the ODE

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{A}\big(\boldsymbol{x}(t),t\big) \tag{2.158}$$

with boundary condition $\boldsymbol{x}(t') = \boldsymbol{x}'$, then by applying the chain rule, we see that

$$P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t') = \delta\big(\boldsymbol{x} - \boldsymbol{x}(t;\boldsymbol{x}')\big) \tag{2.159}$$

solves the Liouville equation. Thus, the probability density remains a δ-function for all time.

2.6.4 Stationary Markov processes and ergodic properties

Stationary Markov processes satisfy a time translation invariance:

$$P(\boldsymbol{x}_1,t_1\,;\,\ldots\,;\,\boldsymbol{x}_N,t_N) = P(\boldsymbol{x}_1,t_1+\tau\,;\,\ldots\,;\,\boldsymbol{x}_N,t_N+\tau)\ . \tag{2.160}$$

This means

$$\begin{aligned} P(\boldsymbol{x},t) &= P(\boldsymbol{x}) \\ P(\boldsymbol{x}_1,t_1\,|\,\boldsymbol{x}_2,t_2) &= P(\boldsymbol{x}_1,t_1-t_2\,|\,\boldsymbol{x}_2,0)\ . \end{aligned} \tag{2.161}$$

Consider the case of one space dimension and define the time average

$$\overline{X}_T \equiv \frac{1}{T}\int\limits_{T/2}^{T/2} dt\,x(t)\ . \tag{2.162}$$

We use a bar to denote time averages and angular brackets $\langle\cdots\rangle$ to denote averages over the randomness. Thus, $\langle\overline{X}_T\rangle = \langle x\rangle$, which is time-independent for a stationary Markov process. The variance of $\overline{X}_T$ is

$$\mathsf{Var}\big(\overline{X}_T\big) = \frac{1}{T^2}\int\limits_{T/2}^{T/2}dt\int\limits_{T/2}^{T/2}dt'\,\big\langle x(t)\,x(t')\big\rangle_{\rm c}\ , \tag{2.163}$$

where the connected average is $\langle AB\rangle_{\rm c} = \langle AB\rangle - \langle A\rangle\langle B\rangle$. We define

$$C(t_1 - t_2) \equiv \big\langle x(t_1)\, x(t_2)\big\rangle = \int\limits_{-\infty}^{\infty} dx_1 \int\limits_{-\infty}^{\infty} dx_2\; x_1\, x_2\; P(x_1, t_1\,;\, x_2, t_2)\;. \tag{2.164}$$

If $C(\tau)$ decays to zero sufficiently rapidly with τ, for example as an exponential $e^{-\gamma\tau}$, then $\mathsf{Var}\big(\overline{X}_T\big) \to 0$ as $T \to \infty$, which means that $X_{T\to\infty} = \langle x\rangle$. Thus the time average is the ensemble average, which means the process is *ergodic*.

Wiener-Khinchin theorem

Define the quantity

$$\hat{x}_T(\omega) = \int\limits_{T/2}^{T/2} dt\; x(t)\, e^{i\omega t}\;. \tag{2.165}$$

The *spectral function* $S_T(\omega)$ is given by

$$S_T(\omega) = \Big\langle \frac{1}{T}\,\big|\hat{x}_T(\omega)\big|^2 \Big\rangle\;. \tag{2.166}$$

We are interested in the limit $T \to \infty$. Does $S(\omega) \equiv S_{T\to\infty}(\omega)$ exist?

Observe that

$$\begin{aligned}\Big\langle \big|\hat{x}_T(\omega)\big|^2 \Big\rangle &= \int\limits_{T/2}^{T/2} dt_1 \int\limits_{T/2}^{T/2} dt_2\; e^{i\omega(t_2-t_1)}\; \overbrace{\big\langle x(t_1)\, x(t_2)\big\rangle}^{C(t_1-t_2)} \\ &= \int\limits_{-T}^{T} d\tau\; e^{-i\omega\tau}\, C(\tau)\,\big(T - |\tau|\big)\;.\end{aligned} \tag{2.167}$$

Thus,

$$S(\omega) = \lim_{T\to\infty} \int\limits_{-\infty}^{\infty} d\tau\; e^{-i\omega\tau}\, C(\tau)\left(1 - \frac{|\tau|}{T}\right)\Theta\big(T - |\tau|\big) = \int\limits_{-\infty}^{\infty} d\tau\; e^{-i\omega\tau}\, C(\tau)\;. \tag{2.168}$$

The second equality above follows from Lebesgue's dominated convergence theorem, which you can look up on Wikipedia[11]. We therefore conclude the limit exists and is given by the Fourier transform of the correlation function $C(\tau) = \big\langle x(t)\, x(t+\tau)\big\rangle$.

2.6.5 Approach to stationary solution

We have seen, for example, how in general an arbitrary initial state of the Master equation will converge exponentially to an equilibrium distribution. For stationary Markov processes, the conditional distribution $P(\boldsymbol{x}, t\,|\,\boldsymbol{x}', t')$ converges to an equilibrium distribution $P_{\rm eq}(\boldsymbol{x})$ as $t - t' \to \infty$. How can we understand this convergence in terms of the differential Chapman-Kolmogorov equation? We summarize here the results in §3.7.3 of Gardiner.

[11] If we define the one parameter family of functions $C_T(\tau) = C(\tau)\left(1 - \frac{|\tau|}{T}\right)\Theta(T - |\tau|)$, then as $T \to \infty$ the function $C_T(\tau)\, e^{-i\omega\tau}$ converges pointwise to $C(\tau)\, e^{-i\omega\tau}$, and if $|C(\tau)|$ is integrable on $\mathbb{R}$, the theorem guarantees the second equality in Eqn. 2.168.

Suppose $P_1(\boldsymbol{x},t)$ and $P_2(\boldsymbol{x},t)$ are each solutions to the DCK+ equation, and furthermore that $W(\boldsymbol{x}\,|\,\boldsymbol{x}',t)$, $A_\mu(\boldsymbol{x},t)$, and $B_{\mu\nu}(\boldsymbol{x},t)$ are all independent of t. Define the *Lyapunov functional*

$$K[P_1,P_2,t] = \int d\boldsymbol{x}\,\Big(P_1\ln(P_1/P_2) + P_2 - P_1\Big)\,. \tag{2.169}$$

Since $P_{1,2}(\boldsymbol{x},t)$ are both normalized, the integrals of the last two terms inside the big round brackets cancel. Nevertheless, it is helpful to express K in this way since, factoring out P_1 from the terms inside the brackets, we may use $f(z) = z - \ln z - 1 \geq 0$ for $z \in \mathbb{R}_+$, where $z = P_2/P_1$. Thus, $K \geq 0$, and the minimum value is obtained for $P_1(x,t) = P_2(x,t)$.

Next, evaluate the time derivative $\dot{K}$:

$$\frac{dK}{dt} = \int d\boldsymbol{x}\left\{\frac{\partial P_1}{\partial t}\cdot\Big[\ln P_1 - \ln P_2 + 1\Big] - \frac{\partial P_2}{\partial t}\cdot\frac{P_1}{P_2}\right\}. \tag{2.170}$$

We now use DCK+ to obtain $\partial_t P_{1,2}$ and evaluate the contributions due to drift, diffusion, and jump processes. One finds

$$\left(\frac{dK}{dt}\right)_{\rm drift} = -\sum_\mu \int d\boldsymbol{x}\,\frac{\partial}{\partial x_\mu}\Big[A_\mu\,P_1\,\ln\big(P_1/P_2\big)\Big] \tag{2.171}$$

$$\left(\frac{dK}{dt}\right)_{\rm diff} = -\frac{1}{2}\sum_{\mu,\nu}\int d\boldsymbol{x}\,B_{\mu\nu}\,\frac{\partial\ln(P_1/P_2)}{\partial x_\mu}\,\frac{\partial\ln(P_1/P_2)}{\partial x_\nu} + \frac{1}{2}\int d\boldsymbol{x}\,\frac{\partial^2}{\partial x_\mu\,\partial x_\nu}\Big[B_{\mu\nu}\,P_1\,\ln(P_1/P_2)\Big] \tag{2.172}$$

$$\left(\frac{dK}{dt}\right)_{\rm jump} = \int d\boldsymbol{x}\int d\boldsymbol{x}'\,W(\boldsymbol{x}\,|\,\boldsymbol{x}')\,P_2(\boldsymbol{x}',t)\,\Big[\phi'\ln(\phi/\phi') - \phi + \phi'\Big]\,, \tag{2.173}$$

where $\phi(\boldsymbol{x},t) \equiv P_1(\boldsymbol{x},t)/P_2(\boldsymbol{x},t)$ in the last line. Dropping the total derivative terms, which we may set to zero at spatial infinity, we see that $\dot{K}_{\rm drift} = 0$, $\dot{K}_{\rm diff} \leq 0$, and $\dot{K}_{\rm jump} \leq 0$. Barring pathological cases[12], one has that $K(t)$ is a nonnegative decreasing function. Since $K = 0$ when $P_1(\boldsymbol{x},t) = P_2(\boldsymbol{x},t) = P_{\rm eq}(\boldsymbol{x})$, we see that the Lyapunov analysis confirms that K is strictly decreasing. If we set $P_2(\boldsymbol{x},t) = P_{\rm eq}(\boldsymbol{x})$, we conclude that $P_1(\boldsymbol{x},t)$ converges to $P_{\rm eq}(\boldsymbol{x})$ as $t \to \infty$.

2.7 Appendix : Nonlinear diffusion

2.7.1 PDEs with infinite propagation speed

Starting from an initial probability density $P(\boldsymbol{x},t=0) = \delta(\boldsymbol{x})$, we saw how Fickian diffusion, described by the equation $\partial_t P = \boldsymbol{\nabla}\cdot(D\,\boldsymbol{\nabla}P)$, gives rise to the solution

$$P(\boldsymbol{x},t) = (4\pi Dt)^{-d/2}\,e^{-\boldsymbol{x}^2/4Dt}\,, \tag{2.174}$$

for all $t > 0$, assuming D is a constant. As remarked in §2.2.1, this violates any physical limits on the speed of particle propagation, including that set by special relativity, because $P(\boldsymbol{x},t) > 0$ for all $\boldsymbol{x}$ at any finite value of t.

It's perhaps good to step back at this point and recall the solution to the one-dimensional discrete random walk, where after each time increment the walker moves to the right ($\Delta X = 1$) with probability p and to the left ($\Delta X =$

[12] See Gardiner, §3.7.3.

-1) with probability $1-p$. To make things even simpler we'll consider the case with no drift, *i.e.* $p = \frac{1}{2}$. The distribution for X after N time steps is of the binomial form:

$$P_N(X) = 2^{-N} \binom{N}{\frac{1}{2}(N-X)} . \tag{2.175}$$

Invoking Stirling's asymptotic result $\ln K! = K \ln K - K + \mathcal{O}(\ln K)$ for $K \gg 1$, one has[13]

$$P_N(X) \simeq \sqrt{\frac{2}{\pi N}}\, e^{-X^2/2N} . \tag{2.176}$$

We note that the distribution in Eqn. 2.175 is cut off at $|X| = N$, so that $P_N(X) = 0$ for $|X| > N$. This reflects the fact that the walker travels at a fixed speed of one step per time interval. This feature is lost in Eqn. 2.176, because the approximation which led to this result is not valid in the tails of the distribution. One might wonder about the results of §2.3 in this context, since we ultimately obtained a diffusion form for $P(x,t)$ using an exact functional averaging method. However, since we assumed a Gaussian probability functional for the random forcing $\eta(t)$, there is a finite probability for arbitrarily large values of the forcing. For example, consider the distribution of the integrated force $\phi = \int_{t_1}^{t_2} dt\, \eta(t)$:

$$P(\phi, \Delta t) = \left\langle \delta\!\left(\phi - \int_{t_1}^{t_2} dt\, \eta(t)\right)\right\rangle = \frac{1}{\sqrt{2\pi\Gamma\,\Delta t}}\, e^{-\phi^2/2\Gamma\Delta t} , \tag{2.177}$$

where $\Delta t = t_2 - t_1$. This distribution is nonzero for arbitrarily large values of ϕ.

Mathematically, the diffusion equation is an example of what is known as a *parabolic* partial differential equation. The Navier-Stokes equations of hydrodynamics are also parabolic PDEs. The other two classes are called *elliptical* and *hyperbolic*. Paradigmatic examples of these classes include Laplace's equation (elliptical) and the Helmholtz equation (hyperbolic). Hyperbolic equations propagate information at finite propagation speed. For second order PDEs of the form

$$A_{ij}\, \frac{\partial^2 \Psi}{\partial x_i\, \partial x_j} + B_i\, \frac{\partial \Psi}{\partial x_i} + C\Psi = S \quad , \tag{2.178}$$

the PDE is elliptic if the matrix A is positive definite or negative definite, parabolic if A has one zero eigenvalue, and hyperbolic if A is nondegenerate and indefinite (*i.e.* one positive and one negative eigenvalue). Accordingly, one way to remedy the unphysical propagation speed in the diffusion equation is to deform it to a hyperbolic PDE such as the *telegrapher's equation*,

$$\tau\, \frac{\partial^2 \Psi}{\partial t^2} + \frac{\partial \Psi}{\partial t} + \gamma \Psi = D\, \frac{\partial^2 \Psi}{\partial x^2} . \tag{2.179}$$

When $\gamma = 0$, the solution for the initial condition $\Psi(x,0) = \delta(x)$ is

$$\Psi(x,t) = \frac{1}{\sqrt{4Dt}}\, e^{-t/2\tau}\, I_0\left(\sqrt{\left(\frac{t}{2\tau}\right)^2 - \frac{x^2}{4D\tau}}\right) \Theta\big(\sqrt{D/\tau}\, t - |x|\big) . \tag{2.180}$$

Note that $\Psi(x,t)$ vanishes for $|x| > ct$, where $c = \sqrt{D/\tau}$ is the maximum propagation speed. One can check that in the limit $\tau \to 0$ one recovers the familiar diffusion kernel.

[13] The prefactor in this equation seems to be twice the expected $(2\pi N)^{-1/2}$, but since each step results in $\Delta X = \pm 1$, if we start from $X_0 = 0$ then after N steps X will be even if N is even and odd if N is odd. Therefore the continuum limit for the normalization condition on $P_N(X)$ is $\sum_X P_N(X) \approx \frac{1}{2}\int_{-\infty}^{\infty} dX\, P_N(X) = 1$.

The telegrapher's equation

To derive the telegrapher's equation, consider the section of a transmission line shown in Fig. 2.4. Let $V(x,t)$ be the electrical potential on the top line, with $V = 0$ on the bottom (*i.e.* ground). Per unit length a, the potential drop along the top line is $\Delta V = a\,\partial_x V = -IR - L\,\partial_t I$, and the current drop is $\Delta I = a\,\partial_x I = -GV - C\,\partial_t V$. Differentiating the first equation with respect to x and using the second for $\partial_x I$, one arrives at Eqn. 2.179 with $\tau = LC/(RC+GL)$, $\gamma = RG/(RC+GL)$, and $D = a^2/(RC+GL)$.

2.7.2 The porous medium and p-Laplacian equations

Another way to remedy this problem with the diffusion equation is to consider some nonlinear extensions thereof[14]. Two such examples have been popular in the mathematical literature, the *porous medium equation* (PME),

$$\frac{\partial u}{\partial t} = \nabla^2\big(u^m\big)\ , \tag{2.181}$$

and the *p-Laplacian equation*,

$$\frac{\partial u}{\partial t} = \boldsymbol{\nabla}\cdot\Big(|\boldsymbol{\nabla} u|^{p-2}\,\boldsymbol{\nabla} u\Big)\ . \tag{2.182}$$

Both these equations introduce a nonlinearity whereby the diffusion constant D depends on the field u. For example, the PME can be rewritten $\partial_t u = \boldsymbol{\nabla}\cdot\big(m\,u^{m-1}\boldsymbol{\nabla} u\big)$, whence $D = m\,u^{m-1}$. For the p-Laplacian equation, $D = |\boldsymbol{\nabla} u|^{p-2}$. These nonlinearities strangle the diffusion when u or $|\boldsymbol{\nabla} u|$ gets small, preventing the solution from advancing infinitely fast.

As its name betokens, the PME describes fluid flow in a porous medium. A fluid moving through a porous medium is described by three fundamental equations:

(i) *Continuity:* In a medium with porosity ε, the continuity equation becomes $\varepsilon\,\partial_t\varrho + \nabla\cdot(\varrho\boldsymbol{v}) = 0$, where ϱ is the fluid density. This is because in a volume Ω where the fluid density is changing at a rate $\partial_t\varrho$, the rate of change of fluid mass is $\varepsilon\Omega\,\partial_t\varrho$.

(ii) *Darcy's law:* First articulated in 1856 by the French hydrologist Henry Darcy, this says that the flow velocity is directly proportional to the pressure gradient according to the relation $\boldsymbol{v} = -(K/\mu)\boldsymbol{\nabla} p$, where the *permeability* K depends on the medium but not on the fluid, and μ is the shear viscosity of the fluid.

(iii) *Fluid equation of state:* This is a relation between the pressure p and the density ϱ of the fluid. For ideal gases, $p = A\,\varrho^\gamma$ where A is a constant and $\gamma = c_p/c_V$ is the specific heat ratio.

[14] See J. L. Vazquez, *The Porous Medium Equation* (Oxford, 2006).

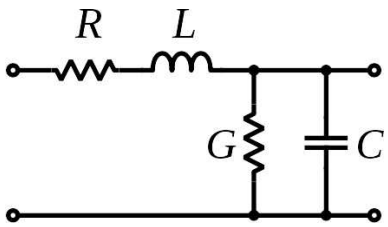

Figure 2.4: Repeating unit of a transmission line. *Credit: Wikipedia*

Putting these three equations together, we obtain

$$\frac{\partial \varrho}{\partial t} = C\,\nabla^2\big(\varrho^m\big)\ , \tag{2.183}$$

where $C = A\gamma k/(k+1)\varepsilon\mu$ and $m = 1+\gamma$.

2.7.3 Illustrative solutions

A class of solution to the PME was discussed in the Russian literature in the early 1950's in a series of papers by Zeldovich, Kompaneets, and Barenblatt. The ZKB solution, which is isotropic in d space dimensions, is of the scaling form,

$$U(r,t) = t^{-\alpha}\,F\big(r\,t^{-\alpha/d}\big) \qquad ; \qquad F(\xi) = \Big(C - k\,\xi^2\Big)_+^{\frac{1}{m-1}} \quad , \tag{2.184}$$

where $r = |\boldsymbol{x}|$,

$$\alpha = \frac{d}{(m-1)d+2} \qquad , \qquad k = \frac{m-1}{2m}\cdot\frac{1}{(m-1)d+2} \quad , \tag{2.185}$$

and the $+$ subscript in the definition of $F(\xi)$ in Eqn. 2.184 indicates that the function is cut off and vanishes when the quantity inside the round brackets becomes negative. We also take $m > 1$, which means that $\alpha < \frac{1}{2}d$. The quantity C is determined by initial conditions. The scaling form is motivated by the fact that the PME conserves the integral of $u(\boldsymbol{x},t)$ over all space, provided the current $\boldsymbol{j} = -mu^{m-1}\boldsymbol{\nabla}u$ vanishes at spatial infinity. Explicitly, we have

$$\int d^dx\,U(\boldsymbol{x},t) = \Omega_d\int_0^\infty dr\,r^{d-1}\,t^{-\alpha}\,F\big(r\,t^{-\alpha/d}\big) = \Omega_d\int_0^\infty ds\,s^{d-1}\,F(s)\ , \tag{2.186}$$

where Ω_d is the total solid angle in d space dimensions. The above integral is therefore independent of t, which means that the integral of U is conserved. Therefore as $t\to 0$, we must have $U(\boldsymbol{x}, t=0) = A\,\delta(\boldsymbol{x})$, where A is a constant which can be expressed in terms of C, m, and d. We plot the behavior of this solution for the case $m = 2$ and $d = 1$ in Fig. 2.5, and compare and contrast it to the solution of the diffusion equation. Note that the solutions to the PME have compact support, *i.e.* they vanish identically for $r > \sqrt{C/k}\,t^{\alpha/d}$, which is consistent with a finite maximum speed of propagation. A similar point source solution to the p-Laplacian equation in $d = 1$ was obtained by Barenblatt:

$$U(x,t) = t^{-m}\Big(C - k\,|\xi|^{1+m^{-1}}\Big)^{\frac{m}{m-1}}\ , \tag{2.187}$$

for arbitrary $C > 0$, with $\xi = x\,t^{-1/2m}$, and $k = (m-1)(2m)^{-(m+1)/m}$.

To derive the ZKB solution of the porous medium equation, it is useful to write the PME in terms of the 'pressure' variable $v = \frac{m}{m-1}\,u^{m-1}$. The PME then takes the form

$$\frac{\partial v}{\partial t} = (m-1)\,v\,\nabla^2 v + (\boldsymbol{\nabla}v)^2\ . \tag{2.188}$$

We seek an isotropic solution in d space dimensions, and posit the scaling form

$$V(\boldsymbol{x},t) = t^{-\lambda}\,G\big(r\,t^{-\mu}\big)\ , \tag{2.189}$$

where $r = |\boldsymbol{x}|$. Acting on isotropic functions, the Laplacian is given by $\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{d-1}{r}\frac{\partial}{\partial r}$. Defining $\xi = r\,t^{-\mu}$, we have

$$\frac{\partial V}{\partial t} = -t^{-1}\Big[\lambda\,G + \mu\,\xi\,G'\Big] \qquad , \qquad \frac{\partial V}{\partial r} = t^{-(\lambda+\mu)}\,G' \qquad , \qquad \frac{\partial^2 V}{\partial r^2} = t^{-(\lambda+2\mu)}\,G'' \quad , \tag{2.190}$$

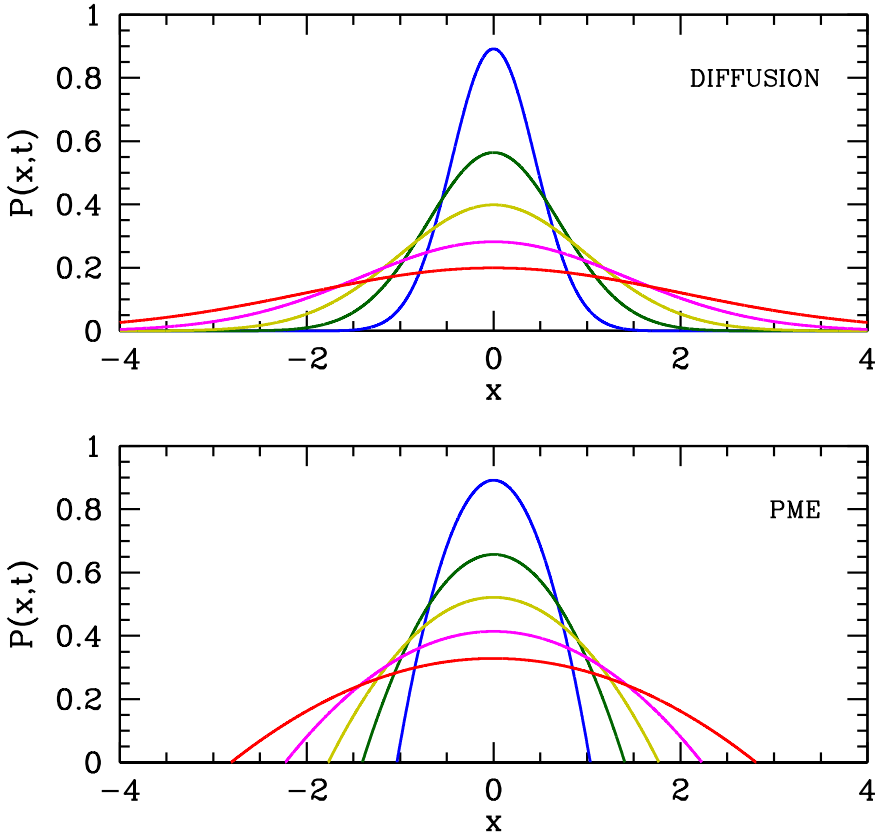

Figure 2.5: Top panel: evolution of the diffusion equation with $D = 1$ and $\sigma = 1$ for times $t = 0.1, 0.25, 0.5, 1.0$, and 2.0. Bottom panel: evolution of the porous medium equation with $m = 2$ and $d = 1$ and C chosen so that $P(x = 0, t = 0.1)$ is equal to the corresponding value in the top panel (*i.e.* the peak of the blue curve).

whence

$$-\Big[\lambda\, G + \mu\, \xi\, G'\Big]\, t^{-1} = \Big[(m-1)\, G\, G'' + (m-1)\, (d-1)\, \xi^{-1}\, G\, G' + (G')^2\Big]\, t^{-2(\lambda+\mu)} \ . \tag{2.191}$$

At this point we can read off the result $\lambda + \mu = \frac{1}{2}$ and eliminate the t variable, which validates our initial scaling form hypothesis. What remains is

$$\lambda\, G + \mu\, \xi G' + (m-1)\, G\, G'' + (m-1)(d-1)\, \xi^{-1}\, G\, G' + (G')^2 = 0 \ . \tag{2.192}$$

Inspection now shows that this equation has a solution of the form $G(\xi) = A - b\,\xi^2$. Plugging this in, we find

$$\lambda = (m-1)\,\alpha \quad , \quad \mu = \frac{\alpha}{d} \quad , \quad b = \frac{\alpha}{2d} \quad , \quad \alpha \equiv \frac{d}{(m-1)\,d+2} \ . \tag{2.193}$$

The quadratic function $G(\xi) = A - b\,\xi^2$ goes negative for $\xi^2 > A/b$, which is clearly unphysical in the context of diffusion. To remedy this, Zeldovich *et al.* proposed to take the maximum value of $G(\xi)$ and zero. Clearly $G = 0$ is a solution, hence $G(\xi) = \big(A - b\,\xi^2\big)_+$ is a solution for $|\xi| < \sqrt{A/b}$ and for $|\xi| > \sqrt{A/b}$, but what about the points $\xi = \pm\sqrt{A/b}$? The concern is that the second derivative $G''(\xi)$ has a delta function singularity at those points, owing to the discontinuity of $G'(\xi)$. However, an examination of Eqn. 2.192 shows that G'' is multiplied by G, and we know that $\lim_{x\to 0} x\,\delta(x) = 0$. The remaining nonzero terms in this equation are then $\big[\mu\,\xi + G'(\xi)\big]G'(\xi)$,

which agreeably vanishes. So we have a solution of the form[15]

$$V(\boldsymbol{x},t) = \frac{1}{t}\left(A'\,t^{2\alpha/d} - \alpha\,\boldsymbol{x}^2\right)_+ \quad , \tag{2.194}$$

where $A' = 2dA$.

2.8 Appendix : Langevin equation for a particle in a harmonic well

Consider next the equation

$$\ddot{X} + \gamma\dot{X} + \omega_0^2 X = \frac{F}{M} + \eta(t)\;, \tag{2.195}$$

where F is a constant force. We write $X = x_0 + x$ and measure x relative to the potential minimum $x_0 = F/M\omega_0^2$, yielding

$$\ddot{x} + \gamma\,\dot{x} + \omega_0^2\,x = \eta(t)\;. \tag{2.196}$$

We solve via Laplace transform. Recall

$$\begin{aligned}\check{x}(z) &= \int\limits_0^\infty dt\,e^{-zt}\,x(t)\\ x(t) &= \int\limits_{\mathcal{C}}\frac{dz}{2\pi i}\,e^{+zt}\,\check{x}(z)\;,\end{aligned} \tag{2.197}$$

where the contour $\mathcal{C}$ proceeds from $c - i\infty$ to $c + i\infty$ such that all poles of the integrand lie to the left of $\mathcal{C}$. Then

$$\begin{aligned}\int\limits_0^\infty dt\,e^{-zt}\left(\ddot{x} + \gamma\,\dot{x} + \omega_0^2\,x\right) &= -(z+\gamma)\,x(0) - \dot{x}(0) + \left(z^2 + \gamma z + \omega_0^2\right)\check{x}(z)\\ &= \int\limits_0^\infty dt\,e^{-zt}\,\eta(t) = \check{\eta}(z)\;.\end{aligned} \tag{2.198}$$

Thus, we have

$$\check{x}(z) = \frac{(z+\gamma)\,x(0) + \dot{x}(0)}{z^2 + \gamma z + \omega_0^2} + \frac{1}{z^2 + \gamma z + \omega_0^2}\int\limits_0^\infty dt\,e^{-zt}\,\eta(t)\;. \tag{2.199}$$

Now we may write

$$z^2 + \gamma z + \omega_0^2 = (z - z_+)(z - z_-)\;, \tag{2.200}$$

where $z_\pm = -\frac{1}{2}\gamma \pm \sqrt{\frac{1}{4}\gamma^2 - \omega_0^2}$. Note that $\mathrm{Re}\,(z_\pm) \le 0$ and that $z_\mp = -\gamma - z_\pm$.

Performing the inverse Laplace transform, we obtain

$$x(t) = \frac{x(0)}{z_+ - z_-}\left(z_+\,e^{z_-t} - z_-\,e^{z_+t}\right) + \frac{\dot{x}(0)}{z_+ - z_-}\left(e^{z_+t} - e^{z_-t}\right) + \int\limits_0^\infty ds\,K(t-s)\,\eta(s) \quad , \tag{2.201}$$

[15] Actually the result $\lim_{x\to 0} x\,\delta(x) = 0$ is valid in the distribution sense, *i.e.* underneath an integral, provided $x\,\delta(x)$ is multiplied by a nonsingular function of x. Thus, Eqn. 2.194 constitutes a *weak solution* to the pressure form of the porous medium equation 2.188. Zeldovich *et al.* found numerically that cutting off the negative part of $A - b\,\xi^2$ is appropriate. Mathematically, Vazquez has shown that when the initial data are taken within a suitable class of integrable functions, the weak solution exists and is unique.

where

$$K(t-s) = \frac{\Theta(t-s)}{(z_+ - z_-)}\left(e^{z_+(t-s)} - e^{z_-(t-s)}\right) \tag{2.202}$$

is the *response kernel* and $\Theta(t-s)$ is the step function which is unity for $t > s$ and zero otherwise. The response is *causal*, *i.e.* $x(t)$ depends on $\eta(s)$ for all previous times $s < t$, but not for future times $s > t$. Note that $K(\tau)$ decays exponentially for $\tau \to \infty$, if $\mathrm{Re}(z_\pm) < 0$. The marginal case where $\omega_0 = 0$ and $z_+ = 0$ corresponds to the diffusion calculation we performed in the previous section.

It is now easy to compute

$$\big\langle x^2(t)\big\rangle_{\rm c} = \Gamma \int_0^t ds\; K^2(s) = \frac{\Gamma}{2\omega_0^2\gamma} \quad (t \to \infty) \tag{2.203}$$

$$\big\langle \dot{x}^2(t)\big\rangle_{\rm c} = \Gamma \int_0^t ds\; \dot{K}^2(s) = \frac{\Gamma}{2\gamma} \qquad (t \to \infty)\;, \tag{2.204}$$

where the connected average is defined by $\langle AB\rangle_{\rm c} = \langle AB\rangle - \langle A\rangle\langle B\rangle$. Therefore,

$$\left\langle \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}M\omega_0^2 x^2 \right\rangle_{t\to\infty} = \frac{M\Gamma}{2\gamma}\;. \tag{2.205}$$

Setting this equal to $2 \times \frac{1}{2}k_{\rm B}T$ by equipartition again yields $\Gamma = 2\gamma k_{\rm B}T/M$.

2.9 Appendix : General Linear Autonomous Inhomogeneous ODEs

2.9.1 Solution by Fourier transform

We can also solve general autonomous linear inhomogeneous ODEs of the form

$$\frac{d^n x}{dt^n} + a_{n-1}\,\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\,\frac{dx}{dt} + a_0\,x = \xi(t)\;. \tag{2.206}$$

We can write this as

$$\mathcal{L}_t\,x(t) = \xi(t)\;, \tag{2.207}$$

where $\mathcal{L}_t$ is the $n^{\rm th}$ order differential operator

$$\mathcal{L}_t = \frac{d^n}{dt^n} + a_{n-1}\,\frac{d^{n-1}}{dt^{n-1}} + \ldots + a_1\,\frac{d}{dt} + a_0\;. \tag{2.208}$$

The general solution to the inhomogeneous equation is given by

$$x(t) = x_{\rm h}(t) + \int_{-\infty}^{\infty} dt'\; G(t,t')\,\xi(t')\;, \tag{2.209}$$

where $G(t,t')$ is the Green's function. Note that $\mathcal{L}_t\,x_{\rm h}(t) = 0$. Thus, in order for eqns. 2.207 and 2.209 to be true, we must have

$$\mathcal{L}_t\,x(t) = \overbrace{\mathcal{L}_t\,x_{\rm h}(t)}^{\text{this vanishes}} + \int_{-\infty}^{\infty} dt'\;\mathcal{L}_t\,G(t,t')\,\xi(t') = \xi(t)\;, \tag{2.210}$$

which means that

$$\mathcal{L}_t\, G(t,t') = \delta(t-t')\ , \tag{2.211}$$

where $\delta(t-t')$ is the Dirac δ-function.

If the differential equation $\mathcal{L}_t\, x(t) = \xi(t)$ is defined over some finite or semi-infinite t interval with prescribed boundary conditions on $x(t)$ at the endpoints, then $G(t,t')$ will depend on t and t' separately. For the case we are now considering, let the interval be the entire real line $t \in (-\infty,\infty)$. Then $G(t,t') = G(t-t')$ is a function of the single variable $t-t'$.

Note that $\mathcal{L}_t = \mathcal{L}\big(\frac{d}{dt}\big)$ may be considered a function of the differential operator $\frac{d}{dt}$. If we now Fourier transform the equation $\mathcal{L}_t\, x(t) = \xi(t)$, we obtain

$$\begin{aligned}\int\limits_{-\infty}^{\infty} dt\, e^{i\omega t}\, \xi(t) &= \int\limits_{-\infty}^{\infty} dt\, e^{i\omega t} \left\{ \frac{d^n}{dt^n} + a_{n-1}\, \frac{d^{n-1}}{dt^{n-1}} + \ldots + a_1\, \frac{d}{dt} + a_0 \right\} x(t) \\ &= \int\limits_{-\infty}^{\infty} dt\, e^{i\omega t} \left\{ (-i\omega)^n + a_{n-1}\, (-i\omega)^{n-1} + \ldots + a_1\, (-i\omega) + a_0 \right\} x(t)\ .\end{aligned} \tag{2.212}$$

Thus, if we define

$$\hat{\mathcal{L}}(\omega) = \sum_{k=0}^{n} a_k\, (-i\omega)^k\ , \tag{2.213}$$

then we have $\hat{\mathcal{L}}(\omega)\, \hat{x}(\omega) = \hat{\xi}(\omega)$, where $a_n \equiv 1$. According to the Fundamental Theorem of Algebra, the $n^{\rm th}$ degree polynomial $\hat{\mathcal{L}}(\omega)$ may be uniquely factored over the complex ω plane into a product over n roots:

$$\hat{\mathcal{L}}(\omega) = (-i)^n\, (\omega-\omega_1)(\omega-\omega_2)\cdots(\omega-\omega_n)\ . \tag{2.214}$$

If the $\{a_k\}$ are all real, then $\big[\hat{\mathcal{L}}(\omega)\big]^* = \hat{\mathcal{L}}(-\omega^*)$, hence if Ω is a root then so is $-\Omega^*$. Thus, the roots appear in pairs which are symmetric about the imaginary axis. *I.e.* if $\Omega = a + ib$ is a root, then so is $-\Omega^* = -a + ib$.

The general solution to the homogeneous equation is

$$x_{\rm h}(t) = \sum_{\sigma=1}^{n} A_\sigma\, e^{-i\omega_\sigma t}\ , \tag{2.215}$$

which involves n arbitrary complex constants A_i. The susceptibility, or Green's function in Fourier space, $\hat{G}(\omega)$ is then

$$\hat{G}(\omega) = \frac{1}{\hat{\mathcal{L}}(\omega)} = \frac{i^n}{(\omega-\omega_1)(\omega-\omega_2)\cdots(\omega-\omega_n)}\ , \tag{2.216}$$

Note that $\big[\hat{G}(\omega)\big]^* = \hat{G}(-\omega)$, which is equivalent to the statement that $G(t-t')$ is a real function of its argument. The general solution to the inhomogeneous equation is then

$$x(t) = x_{\rm h}(t) + \int\limits_{-\infty}^{\infty} dt'\, G(t-t')\, \xi(t')\ , \tag{2.217}$$

where $x_{\rm h}(t)$ is the solution to the homogeneous equation, *i.e.* with zero forcing, and where

$$\begin{aligned} G(t-t') &= \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, e^{-i\omega(t-t')}\, \hat{G}(\omega) \\ &= i^n \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, \frac{e^{-i\omega(t-t')}}{(\omega-\omega_1)(\omega-\omega_2)\cdots(\omega-\omega_n)} \\ &= \sum_{\sigma=1}^{n} \frac{e^{-i\omega_\sigma(t-t')}}{i\,\hat{\mathcal{L}}'(\omega_\sigma)}\,\Theta(t-t')\;, \end{aligned} \tag{2.218}$$

where we assume that $\mathrm{Im}\,\omega_\sigma < 0$ for all σ. This guarantees *causality* – the *response* $x(t)$ to the *influence* $\xi(t')$ is nonzero only for $t > t'$.

As an example, consider the familiar case

$$\begin{aligned} \hat{\mathcal{L}}(\omega) &= -\omega^2 - i\gamma\omega + \omega_0^2 \\ &= -(\omega-\omega_+)\,(\omega-\omega_-)\;, \end{aligned} \tag{2.219}$$

with $\omega_\pm = -\frac{i}{2}\gamma \pm \beta$, and $\beta = \sqrt{\omega_0^2 - \frac{1}{4}\gamma^2}$. This yields $\hat{\mathcal{L}}'(\omega_\pm) = \mp(\omega_+ - \omega_-) = \mp 2\beta$, hence according to equation 2.218,

$$\begin{aligned} G(s) &= \left\{ \frac{e^{-i\omega_+ s}}{i\mathcal{L}'(\omega_+)} + \frac{e^{-i\omega_- s}}{i\mathcal{L}'(\omega_-)} \right\} \Theta(s) \\ &= \left\{ \frac{e^{-\gamma s/2}\, e^{-i\beta s}}{-2i\beta} + \frac{e^{-\gamma s/2}\, e^{i\beta s}}{2i\beta} \right\} \Theta(s) = \beta^{-1}\, e^{-\gamma s/2}\, \sin(\beta s)\, \Theta(s)\;. \end{aligned} \tag{2.220}$$

Now let us evaluate the two-point correlation function $\big\langle x(t)\, x(t')\big\rangle$, assuming the noise is correlated according to $\big\langle \xi(s)\, \xi(s')\big\rangle = \phi(s-s')$. We assume $t, t' \to \infty$ so the transient contribution $x_{\rm h}$ is negligible. We then have

$$\big\langle x(t)\, x(t')\big\rangle = \int\limits_{-\infty}^{\infty} ds \int\limits_{-\infty}^{\infty} ds'\, G(t-s)\, G(t'-s')\, \big\langle \xi(s)\, \xi(s')\big\rangle = \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, \hat{\phi}(\omega)\, \big|\hat{G}(\omega)\big|^2\, e^{i\omega(t-t')}\;. \tag{2.221}$$

2.9.2 Higher order ODEs

Note that any $n^{\rm th}$ order ODE, of the general form

$$\frac{d^n x}{dt^n} = F\left(x\,,\, \frac{dx}{dt}\,,\, \ldots\,,\, \frac{d^{n-1}x}{dt^{n-1}}\right)\,, \tag{2.222}$$

may be represented by the first order system $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$. To see this, define $\varphi_k = d^{k-1}x/dt^{k-1}$, with $k = 1, \ldots, n$. Thus, for $k < n$ we have $\dot{\varphi}_k = \varphi_{k+1}$, and $\dot{\varphi}_n = F$. In other words,

$$\overbrace{\frac{d}{dt}\begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_{n-1} \\ \varphi_n \end{pmatrix}}^{\dot{\boldsymbol{\varphi}}} = \overbrace{\begin{pmatrix} \varphi_2 \\ \vdots \\ \varphi_n \\ F\big(\varphi_1, \ldots, \varphi_p\big) \end{pmatrix}}^{\boldsymbol{V}(\boldsymbol{\varphi})}\;. \tag{2.223}$$

An inhomogeneous linear $n^{\rm th}$ order ODE,

$$\frac{d^n x}{dt^n} + a_{n-1}\,\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\,\frac{dx}{dt} + a_0\,x = \xi(t) \tag{2.224}$$

may be written in matrix form, as

$$\frac{d}{dt}\begin{pmatrix}\varphi_1\\ \varphi_2\\ \vdots\\ \varphi_n\end{pmatrix} = \overbrace{\begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1}\end{pmatrix}}^{Q}\begin{pmatrix}\varphi_1\\ \varphi_2\\ \vdots\\ \varphi_n\end{pmatrix} + \overbrace{\begin{pmatrix}0\\ 0\\ \vdots\\ \xi(t)\end{pmatrix}}^{\boldsymbol{\xi}}\ . \tag{2.225}$$

Thus,

$$\dot{\boldsymbol{\varphi}} = Q\,\boldsymbol{\varphi} + \boldsymbol{\xi}\ , \tag{2.226}$$

and if the coefficients c_k are time-independent, *i.e.* the ODE is *autonomous*.

For the homogeneous case where $\xi(t) = 0$, the solution is obtained by exponentiating the constant matrix Qt:

$$\boldsymbol{\varphi}(t) = \exp(Qt)\,\boldsymbol{\varphi}(0)\ ; \tag{2.227}$$

the exponential of a matrix may be given meaning by its Taylor series expansion. If the ODE is not autonomous, then $Q = Q(t)$ is time-dependent, and the solution is given by the path-ordered exponential,

$$\boldsymbol{\varphi}(t) = \mathsf{P}\exp\left\{\int\limits_0^t dt'\,Q(t')\right\}\boldsymbol{\varphi}(0)\ , \tag{2.228}$$

where P is the *path ordering operator* which places earlier times to the right. As defined, the equation $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$ is autonomous, since the t-advance mapping g_t depends only on t and on no other time variable. However, by extending the phase space $\mathbb{M} \ni \boldsymbol{\varphi}$ from $\mathbb{M} \to \mathbb{M}\times\mathbb{R}$, which is of dimension $n+1$, one can describe arbitrary time-dependent ODEs.

In general, path ordered exponentials are difficult to compute analytically. We will henceforth consider the autonomous case where Q is a constant matrix in time. We will assume the matrix Q is real, but other than that it has no helpful symmetries. We can however decompose it into left and right eigenvectors:

$$Q_{ij} = \sum_{\sigma=1}^{n} \nu_\sigma\,R_{\sigma,i}\,L_{\sigma,j}\ . \tag{2.229}$$

Or, in bra-ket notation, $Q = \sum_\sigma \nu_\sigma\,|R_\sigma\rangle\langle L_\sigma|$. We adopt the normalization convention $\big\langle\, L_\sigma \,\big|\, R_{\sigma'}\,\big\rangle = \delta_{\sigma\sigma'}$, where $\{\nu_\sigma\}$ are the eigenvalues of Q. The eigenvalues may be real or imaginary. Since the characteristic polynomial $P(\nu) = \det(\nu\,\mathbb{I} - Q)$ has real coefficients, we know that the eigenvalues of Q are either real or come in complex conjugate pairs.

Consider, for example, the $n = 2$ system we studied earlier. Then

$$Q = \begin{pmatrix}0 & 1\\ -\omega_0^2 & -\gamma\end{pmatrix}\ . \tag{2.230}$$

The eigenvalues are as before: $\nu_\pm = -\frac{1}{2}\gamma \pm \sqrt{\frac{1}{4}\gamma^2 - \omega_0^2}$. The left and right eigenvectors are

$$L_\pm = \frac{\pm 1}{\nu_+ - \nu_-}\begin{pmatrix}-\nu_\mp & 1\end{pmatrix} \qquad , \qquad R_\pm = \begin{pmatrix}1\\ \nu_\pm\end{pmatrix}\ . \tag{2.231}$$

The utility of working in a left-right eigenbasis is apparent once we reflect upon the result

$$f(Q) = \sum_{\sigma=1}^{n} f(\nu_\sigma) \, \big| \, R_\sigma \, \big\rangle \big\langle \, L_\sigma \, \big| \tag{2.232}$$

for any function f. Thus, the solution to the general autonomous homogeneous case is

$$\begin{aligned} \big| \, \boldsymbol{\varphi}(t) \, \big\rangle &= \sum_{\sigma=1}^{n} e^{\nu_\sigma t} \, \big| \, R_\sigma \, \big\rangle \big\langle \, L_\sigma \, \big| \, \boldsymbol{\varphi}(0) \, \big\rangle \\ \varphi_i(t) &= \sum_{\sigma=1}^{n} e^{\nu_\sigma t} \, R_{\sigma,i} \sum_{j=1}^{n} L_{\sigma,j} \, \varphi_j(0) \; . \end{aligned} \tag{2.233}$$

If $\mathsf{Re}\,(\nu_\sigma) \le 0$ for all σ, then the initial conditions $\boldsymbol{\varphi}(0)$ are forgotten on time scales $\tau_\sigma = \nu_\sigma^{-1}$. Physicality demands that this is the case.

Now let's consider the inhomogeneous case where $\xi(t) \neq 0$. We begin by recasting eqn. 2.226 in the form

$$\frac{d}{dt}\big(e^{-Qt} \, \boldsymbol{\varphi}\big) = e^{-Qt} \, \boldsymbol{\xi}(t) \; . \tag{2.234}$$

We can integrate this directly:

$$\boldsymbol{\varphi}(t) = e^{Qt} \, \boldsymbol{\varphi}(0) + \int_0^t ds \, e^{Q(t-s)} \, \boldsymbol{\xi}(s) \; . \tag{2.235}$$

In component notation,

$$\varphi_i(t) = \sum_{\sigma=1}^{n} e^{\nu_\sigma t} \, R_{\sigma,i} \, \big\langle \, L_\sigma \, \big| \, \boldsymbol{\varphi}(0) \, \big\rangle + \sum_{\sigma=1}^{n} R_{\sigma,i} \int_0^t ds \, e^{\nu_\sigma (t-s)} \, \big\langle \, L_\sigma \, \big| \, \boldsymbol{\xi}(s) \, \big\rangle . \tag{2.236}$$

Note that the first term on the RHS is the solution to the homogeneous equation, as must be the case when $\boldsymbol{\xi}(s) = 0$.

The solution in eqn. 2.236 holds for general Q and $\boldsymbol{\xi}(s)$. For the particular form of Q and $\xi(s)$ in eqn. 2.225, we can proceed further. For starters, $\langle L_\sigma | \boldsymbol{\xi}(s) \rangle = L_{\sigma,n} \, \xi(s)$. We can further exploit a special feature of the Q matrix to analytically determine all its left and right eigenvectors. Applying Q to the right eigenvector $|R_\sigma\rangle$, we find $R_{\sigma,j} = \nu_\sigma \, R_{\sigma,j-1}$ for $j > 1$. We are free to choose $R_{\sigma,1} = 1$ for all σ and defer the issue of normalization to the derivation of the left eigenvectors. Thus, we obtain the pleasingly simple result, $R_{\sigma,k} = \nu_\sigma^{k-1}$. Applying Q to the left eigenvector $\langle L_\sigma |$, we obtain

$$\begin{aligned} -a_0 \, L_{\sigma,n} &= \nu_\sigma \, L_{\sigma,1} \\ L_{\sigma,j-1} - a_{j-1} \, L_{\sigma,n} &= \nu_\sigma \, L_{\sigma,j} \qquad (j > 1) \; . \end{aligned} \tag{2.237}$$

From these equations we may derive

$$L_{\sigma,k} = -\frac{L_{\sigma,n}}{\nu_\sigma} \sum_{j=0}^{k-1} a_j \, \nu_\sigma^{j-k-1} = \frac{L_{\sigma,n}}{\nu_\sigma} \sum_{j=k}^{n} a_j \, \nu_\sigma^{j-k-1} \; . \tag{2.238}$$

The equality in the above equation is derived using the result $P(\nu_\sigma) = \sum_{j=0}^{n} a_j \, \nu_\sigma^j = 0$. Recall also that $a_n \equiv 1$. We now impose the normalization condition,

$$\sum_{k=1}^{n} L_{\sigma,k} \, R_{\sigma,k} = 1 \; . \tag{2.239}$$

This condition determines our last remaining unknown quantity (for a given σ), $L_{\sigma,p}$:

$$\big\langle\, L_\sigma \,\big|\, R_\sigma \,\big\rangle = L_{\sigma,n} \sum_{k=1}^{n} k\, a_k\, \nu_\sigma^{k-1} = P'(\nu_\sigma)\, L_{\sigma,n}\ , \tag{2.240}$$

where $P'(\nu)$ is the first derivative of the characteristic polynomial. Thus, we find $L_{\sigma,n} = 1/P'(\nu_\sigma)$.

Now let us evaluate the general two-point correlation function,

$$C_{jj'}(t,t') \equiv \big\langle \varphi_j(t)\, \varphi_{j'}(t') \big\rangle - \big\langle \varphi_j(t) \big\rangle \big\langle \varphi_{j'}(t') \big\rangle\ . \tag{2.241}$$

We write

$$\big\langle \xi(s)\, \xi(s') \big\rangle = \phi(s-s') = \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, \hat\phi(\omega)\, e^{-i\omega(s-s')}\ . \tag{2.242}$$

When $\hat\phi(\omega)$ is constant, we have $\big\langle \xi(s)\, \xi(s') \big\rangle = \hat\phi(t)\, \delta(s-s')$. This is the case of so-called *white noise*, when all frequencies contribute equally. The more general case when $\hat\phi(\omega)$ is frequency-dependent is known as *colored noise*. Appealing to eqn. 2.236, we have

$$C_{jj'}(t,t') = \sum_{\sigma,\sigma'} \frac{\nu_\sigma^{j-1}}{P'(\nu_\sigma)}\, \frac{\nu_{\sigma'}^{j'-1}}{P'(\nu_{\sigma'})} \int\limits_0^t ds\, e^{\nu_\sigma (t-s)} \int\limits_0^{t'} ds'\, e^{\nu_{\sigma'}(t'-s')}\, \phi(s-s') \tag{2.243}$$

$$= \sum_{\sigma,\sigma'} \frac{\nu_\sigma^{j-1}}{P'(\nu_\sigma)}\, \frac{\nu_{\sigma'}^{j'-1}}{P'(\nu_{\sigma'})} \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, \frac{\hat\phi(\omega)\, (e^{-i\omega t} - e^{\nu_\sigma t})(e^{i\omega t'} - e^{\nu_{\sigma'} t'})}{(\omega - i\nu_\sigma)(\omega + i\nu_{\sigma'})}\ . \tag{2.244}$$

In the limit $t, t' \to \infty$, assuming $\mathsf{Re}\,(\nu_\sigma) < 0$ for all σ (*i.e.* no diffusion), the exponentials $e^{\nu_\sigma t}$ and $e^{\nu_{\sigma'} t'}$ may be neglected, and we then have

$$C_{jj'}(t,t') = \sum_{\sigma,\sigma'} \frac{\nu_\sigma^{j-1}}{P'(\nu_\sigma)}\, \frac{\nu_{\sigma'}^{j'-1}}{P'(\nu_{\sigma'})} \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, \frac{\hat\phi(\omega)\, e^{-i\omega(t-t')}}{(\omega - i\nu_\sigma)(\omega + i\nu_{\sigma'})}\ . \tag{2.245}$$

2.9.3 Kramers-Krönig relations

Suppose $\hat\chi(\omega) \equiv \hat G(\omega)$ is analytic in the UHP[16]. Then for all ν, we must have

$$\int\limits_{-\infty}^{\infty} \frac{d\nu}{2\pi}\, \frac{\hat\chi(\nu)}{\nu - \omega + i\epsilon} = 0\ , \tag{2.246}$$

where ϵ is a positive infinitesimal. The reason is simple: just close the contour in the UHP, assuming $\hat\chi(\omega)$ vanishes sufficiently rapidly that Jordan's lemma can be applied. Clearly this is an extremely weak restriction on $\hat\chi(\omega)$, given the fact that the denominator already causes the integrand to vanish as $|\omega|^{-1}$.

Let us examine the function

$$\frac{1}{\nu - \omega + i\epsilon} = \frac{\nu - \omega}{(\nu-\omega)^2 + \epsilon^2} - \frac{i\epsilon}{(\nu-\omega)^2 + \epsilon^2}\ . \tag{2.247}$$

[16]In this section, we use the notation $\hat\chi(\omega)$ for the susceptibility, rather than $\hat G(\omega)$

which we have separated into real and imaginary parts. Under an integral sign, the first term, in the limit $\epsilon \to 0$, is equivalent to taking a *principal part* of the integral. That is, for any function $F(\nu)$ which is regular at $\nu = \omega$,

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \frac{d\nu}{2\pi} \, \frac{\nu - \omega}{(\nu - \omega)^2 + \epsilon^2} \, F(\nu) \equiv \wp \int_{-\infty}^{\infty} \frac{d\nu}{2\pi} \, \frac{F(\nu)}{\nu - \omega} \; . \tag{2.248}$$

The *principal part* symbol $\wp$ means that the singularity at $\nu = \omega$ is elided, either by smoothing out the function $1/(\nu - \epsilon)$ as above, or by simply cutting out a region of integration of width ϵ on either side of $\nu = \omega$.

The imaginary part is more interesting. Let us write

$$h(u) \equiv \frac{\epsilon}{u^2 + \epsilon^2} \; . \tag{2.249}$$

For $|u| \gg \epsilon$, $h(u) \simeq \epsilon/u^2$, which vanishes as $\epsilon \to 0$. For $u = 0$, $h(0) = 1/\epsilon$ which diverges as $\epsilon \to 0$. Thus, $h(u)$ has a huge peak at $u = 0$ and rapidly decays to 0 as one moves off the peak in either direction a distance greater that ϵ. Finally, note that

$$\int_{-\infty}^{\infty} du \, h(u) = \pi \; , \tag{2.250}$$

a result which itself is easy to show using contour integration. Putting it all together, this tells us that

$$\lim_{\epsilon \to 0} \, \frac{\epsilon}{u^2 + \epsilon^2} = \pi \delta(u) \; . \tag{2.251}$$

Thus, for positive infinitesimal ϵ,

$$\frac{1}{u \pm i\epsilon} = \frac{\wp}{u} \mp i\pi\delta(u) \; , \tag{2.252}$$

a most useful result.

We now return to our initial result 2.246, and we separate $\hat{\chi}(\omega)$ into real and imaginary parts:

$$\hat{\chi}(\omega) = \hat{\chi}'(\omega) + i\hat{\chi}''(\omega) \; . \tag{2.253}$$

(In this equation, the primes do not indicate differentiation with respect to argument.) We therefore have, for every real value of ω,

$$0 = \int_{-\infty}^{\infty} \frac{d\nu}{2\pi} \Big[\chi'(\nu) + i\chi''(\nu) \Big] \Big[\frac{\wp}{\nu - \omega} - i\pi\delta(\nu - \omega) \Big] \; . \tag{2.254}$$

Taking the real and imaginary parts of this equation, we derive the *Kramers-Krönig relations*:

$$\chi'(\omega) \;\; = \;\; +\wp \int_{-\infty}^{\infty} \frac{d\nu}{\pi} \, \frac{\hat{\chi}''(\nu)}{\nu - \omega} \tag{2.255}$$

$$\chi''(\omega) \;\; = \;\; -\wp \int_{-\infty}^{\infty} \frac{d\nu}{\pi} \, \frac{\hat{\chi}'(\nu)}{\nu - \omega} \; . \tag{2.256}$$

2.10 Appendix : Method of Characteristics

2.10.1 Quasilinear partial differential equations

Consider the quasilinear PDE

$$a_1(\boldsymbol{x},\phi)\,\frac{\partial\phi}{\partial x_1} + a_2(\boldsymbol{x},\phi)\,\frac{\partial\phi}{\partial x_2} + \ldots + a_N(\boldsymbol{x},\phi)\,\frac{\partial\phi}{\partial x_N} = b(\boldsymbol{x},\phi)\ . \tag{2.257}$$

This PDE is called 'quasilinear' because it is linear in the derivatives $\partial\phi/\partial x_j$. The N independent variables are the elements of the vector $\boldsymbol{x} = (x_1,\ldots,x_N)$. A solution is a function $\phi(\boldsymbol{x})$ which satisfies the PDE.

Now consider a curve $\boldsymbol{x}(s)$ parameterized by a single real variable s satisfying

$$\frac{dx_j}{ds} = a_j\big(\boldsymbol{x},\phi(\boldsymbol{x})\big)\ , \tag{2.258}$$

where $\phi(\boldsymbol{x})$ is a solution of eqn. 2.257. Along such a curve, which is called a *characteristic*, the variation of ϕ is

$$\frac{d\phi}{ds} = \sum_{j=1}^{N}\frac{\partial\phi}{\partial x_j}\,\frac{dx_j}{ds} = b\big(\boldsymbol{x}(s),\phi\big)\ . \tag{2.259}$$

Thus, we have converted our PDE into a set of $N+1$ ODEs. To integrate, we must supply some initial conditions of the form

$$g\big(\boldsymbol{x},\phi\big)\Big|_{s=0} = 0\ . \tag{2.260}$$

This defines an $(N-1)$-dimensional hypersurface, parameterized by $\{\zeta_1,\ldots,\zeta_{N-1}\}$:

$$\begin{aligned} x_j(s=0) &= h_j(\zeta_1,\ldots,\zeta_{N-1}) \qquad , \qquad j\in\{1,\ldots,N\} \\ \phi(s=0) &= f(\zeta_1,\ldots,\zeta_{N-1})\ . \end{aligned} \tag{2.261}$$

If we can solve for all the characteristic curves, then the solution of the PDE follows. For every $\boldsymbol{x}$, we identify the characteristic curve upon which $\boldsymbol{x}$ lies. The characteristics are identified by their parameters $(\zeta_1,\ldots,\zeta_{N-1})$. The solution is then $\phi(\boldsymbol{x}) = \phi(s;\zeta_1,\ldots,\zeta_{N-1})$. If two or more characteristics cross, the solution is multi-valued, or a shock has occurred.

2.10.2 Example

Consider the PDE

$$\phi_t + t^2\,\phi_x = -x\,\phi\ . \tag{2.262}$$

We identify $a_1(t,x,\phi) = 1$ and $a_2(t,x,\phi) = t^2$, as well as $b(t,x,\phi) = -x\,\phi$. The characteristics are curves $\big(t(s),x(s)\big)$ satisfing

$$\frac{dt}{ds} = 1 \qquad , \qquad \frac{dx}{ds} = t^2\ . \tag{2.263}$$

The variation of ϕ along each of the characteristics is given by

$$\frac{d\phi}{ds} = -x\,\phi\ . \tag{2.264}$$

The initial data are expressed parametrically as

$$t(s=0)=0 \quad , \quad x(s=0)=\zeta \quad , \quad \phi(s=0)=f(\zeta) \quad . \tag{2.265}$$

We now solve for the characteristics. We have

$$\frac{dt}{ds}=1 \quad \Rightarrow \quad t(s,\zeta)=s\ . \tag{2.266}$$

It then follows that

$$\frac{dx}{ds}=t^2=s^2 \quad \Rightarrow \quad x(s,\zeta)=\zeta+\tfrac{1}{3}s^3\ . \tag{2.267}$$

Finally, we have

$$\frac{d\phi}{ds}=-x\,\phi=-\big(\zeta+\tfrac{1}{3}s^3\big)\,\phi \quad \Rightarrow \quad \phi(s,\zeta)=f(\zeta)\,\exp\big(-\tfrac{1}{12}s^4-s\zeta\big)\ . \tag{2.268}$$

We may now eliminate (ζ,s) in favor of (x,t), writing $s=t$ and $\zeta=x-\frac{1}{3}t^3$, yielding the solution

$$\phi(x,t)=\phi\big(x-\tfrac{1}{3}t^3,t=0\big)\,\exp\big(\tfrac{1}{4}t^4-xt\big)\ . \tag{2.269}$$