Contents of this Document [ntc6]

- Regression theorem for autocorrelation functions [nex39]
- Birth death processes (specifications, models, levels of description) [nln18]
- Birth and death of single species [nln19]
- Birth-death master equation: stationary state [nln17]
- Nonlinear birth-death process

Time-Dependent Probability Distributions [nln50]

Stochastic processes describe systems evolving probabilistically in time. Consider a process characterized by the stochastic variable $X(t)$. In general, the time evolution of $X(t)$ is encoded in a hierarchy of time-dependent joint probability distributions:

$$P(x_1,t_1), \quad P(x_1,t_1;x_2,t_2), \quad \ldots, \quad P(x_1,t_1;x_2,t_2;\ldots;x_n t_n), \quad \ldots$$

with time-ordering $t_1 \geq t_2 \geq \cdots \geq t_n \geq \cdots$ implied.

Attributes:

- $P(x_1,t_1;x_2,t_2;\ldots) \geq 0$,
- $\displaystyle\int dx_1 P(x_1,t_1;x_2,t_2,\ldots) = P(x_2,t_2;\ldots)$,
- $\displaystyle\int dx_1 P(x_1,t_1) = 1$.

Conditional probability distribution: $\displaystyle P(x_1,t_1|x_2,t_2) = \frac{P(x_1,t_1;x_2,t_2)}{P(x_2,t_2)}$.

Attributes: $\displaystyle\int dx_2 P(x_1,t_1|x_2,t_2) = P(x_1,t_1), \quad \int dx_1 P(x_1,t_1|x_2,t_2) = 1$.

More generally: $\displaystyle P(x_1,t_1;\ldots|\bar{x}_1,\bar{t}_1;\ldots) = \frac{P(x_1,t_1;\ldots;\bar{x}_1,\bar{t}_1;\ldots)}{P(\bar{x}_1,\bar{t}_1;\ldots)}$.

Autocorrelation function: $\displaystyle\langle X(t_1)X(t_2)\rangle = \int dx_1 \int dx_2\, x_1 x_2 P(x_1,t_1;x_2,t_2)$.

More generally:

$$\Big\langle [X(t_1)]^{m_1} \cdots [X(t_n)]^{m_n} \Big\rangle = \int dx_1 \cdots \int dx_n\, x_1^{m_1} \cdots x_n^{m_n} P(x_1,t_1;\ldots;x_n,t_n).$$

Stationary processes:

- $P(x_1,t_1;x_2,t_2;\ldots) = P(x_1,t_1+\tau;x_2,t_2+\tau;\ldots)$ for any τ,
- $P(x_1,t_1) = P(x_1,0)$ (time-independent),
- $P(x_1,t_1|x_2,t_2) = P(x_1,t_1-t_2|x_2,0)$ (initial condition),
- $\langle X(t_1)X(t_2)\rangle = \langle X(t_1-t_2)X(0)\rangle$.

Equilibrium implies stationarity but not vice versa.

Degrees of Memory [nln51]

Identification of three types of stochastic processes.

The following time ordering is assumed: $t_1 \geq t_2 \geq \cdots \geq \bar{t}_1 \geq \bar{t}_2 \geq \cdots$.

1. **Future independent of present and past.**
   Completely factorizing process.
   $P(x_1, t_1; x_2, t_2; \ldots | \bar{x}_1, \bar{t}_1; \bar{x}_2, \bar{t}_2, \ldots) = P(x_1, t_1) P(x_2, t_2) \cdots$.

   Example: Gaussian white noise: $P(x, t) = (2\pi\sigma^2)^{-1/2} e^{-x^2/2\sigma^2}$,
   $\langle X(t) \rangle = 0, \quad \langle X(t) X(t') \rangle = \sigma^2 \delta(t - t')$,
   $\int d\tau \langle X(t) X(t + \tau) \rangle e^{i\omega\tau} = \sigma^2 = \text{const} \quad$ (spectral density).

2. **Future dependent on present only.**
   Markov process.
   $P(x_1, t_1; x_2, t_2; \ldots | \bar{x}_1, \bar{t}_1; \bar{x}_2, \bar{t}_2, \ldots) = P(x_1, t_1; x_2, t_2; \ldots | \bar{x}_1, \bar{t}_1)$.

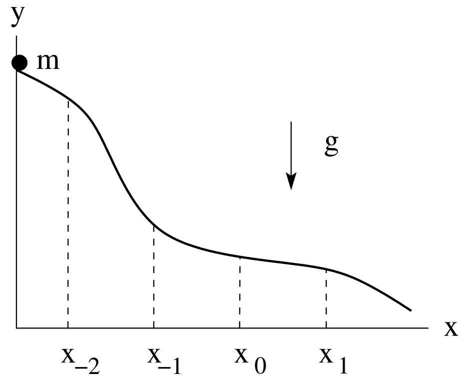
3. **Future dependent on present and past.**
   Non-Markovian process.
   $P(x_1, t_1; x_2, t_2; \ldots | \bar{x}_1, \bar{t}_1; \bar{x}_2, \bar{t}_2, \ldots)$.

Comments:

- Type-2 processes are the main focus in parts 6 and 7 of this course.
- Connections discussed in part 8 of this course: (i) type-1 and type-2 processes interlinked in Langevin equation, (ii) type-2 and type-3 processes interlinked in generalized Langevin equation.
- The same physical process may be described as a type-2 process or a type-3 process depending on the level of description and the choice variables.

Markovian or Non-Markovian I [nln52]

Consider a physical ensemble of particles sliding down some uneven slope, driven by gravity as shown. Each particle starts with random initial velocity at the top of the slope.



We examine probability distributions with a one-component dynamical variable x and probability distributions with a two-component dynamical variable $[x, v]$.

Which of the following probability distributions are broad and which are sharp? Which conditions are redundant?

[1] $P(x_1, t_1)$,

[2] $P(x_1, t_1|x_0, t_0)$,

[3] $P(x_1, t_1|x_0, t_0; x_{-1}, t_{-1})$

[4] $P(x_1, t_1|x_0, t_0; x_{-1}, t_{-1}; x_{-2}, t_{-2})$.

[5] $P([x_1, v_1], t_1)$,

[6] $P([x_1, v_1], t_1|[x_0, v_0], t_0)$,

[7] $P([x_1, v_1], t_1|[x_0, v_0], t_0; [x_{-1}, v_{-1}], t_{-1})$

Answers: [1], [2], [5] are broad. The last condition in [4], [7] is redundant.

Comment: The above answers are independent of whether attenuation is absent or present as long as the motion is deterministic.

Markovian or Non-Markovian II [nln53]

Consider a dilute classical gas, i.e. a physical ensemble of free massive particles moving in a box. The particles move with constant velocity between collisions. The average time between collisons is τ_c. Here we focus on the motion of the particles in x-direction.

Short time intervals: $t \ll \tau_c$ (no collisions during any time interval)

(i) Probability distribution of two-component random variable $[x, \dot{x}]$:

$$P([x_1, \dot{x}_1], t_1 | [x_0, \dot{x}_0], t_0) = \delta\big(\dot{x}_1 - \dot{x}_0\big)\delta\big(x_1 - x_0 - \dot{x}_0(t_1 - t_0)\big).$$

The motion is deterministic. The conditional probability distribution is sharp. The process thus described is Markovian. Any additional condition $[x_{-1}, \dot{x}_{-1}]$ associated with a prior time t_{-1} is redundant.

(ii) Probability distribution of one-component random variable x:

(a) If we insist on a Markovian description, by means of the conditional probability distribution,

$$P_1(x_1, t_1 | x_0, t_0),$$

we obtain a broad distribution even though the process is deterministic.

(b) If we insist on a sharp distribution we must choose a non-Markovian description, by means of the probability distribution with two conditions,

$$P_2(x_1, t_1 | x_0, t_0; x_{-1}, t_{-1}).$$

Any further condition x_{-2} associated with a prior time t_{-2} is again redundant.

The contraction of the level of description from $[x, \dot{x}]$ as in (i) to x as in (ii) of one and the same deterministic time evolution shifts information about the process into memory (cf. [nln15]).

Long time intervals: $t \gg \tau_c$ (many collisions during every time interval)

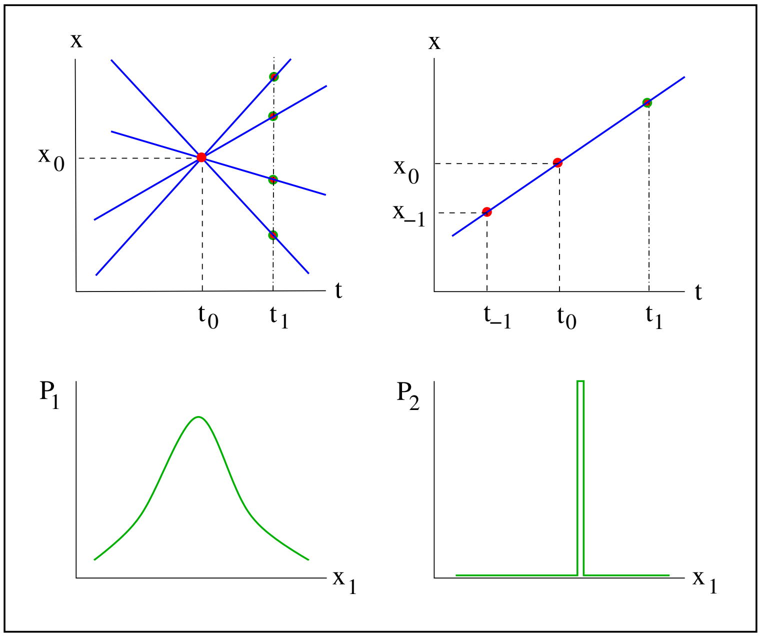
(iii) Markov process $P_1(x_1, t_1 | x_0, t_0)$ is probabilistic (not deterministic).

(iv) Non-Markov process $P_2(x_1, t_1 | x_0, t_0; x_{-1}, t_{-1})$ is also probabilistic.

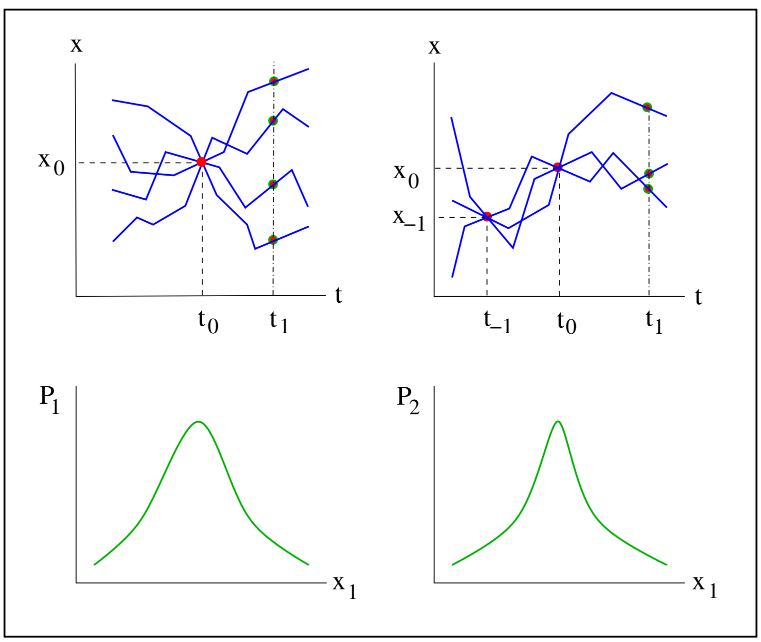
Both conditional probability distributions are broad. The second condition narrows P_2 down relativ to P_1 if $t_0 - t_{-1}$ is short. With increasing $t_0 - t_{-1}$ the effect of the second condition fades away.

[Illustrations on next page]

Short time intervals: $t \ll \tau_c$ (no collisions during any time interval)



Long time intervals: $t \gg \tau_c$ (many collisions during every time interval)



Contraction – memory – time scales [nln15]

microscopic dynamics	**⇒ contraction ⇒**	**stochastic dynamics**
future state determined by present state alone	focus on subset of dynamical variables	future state determined by present and past states
deterministic time evolution of dynamic variables	⇓	ignoring memory of past makes dynamics of selected variables probabilistic
	judicious choice: slow variables and long time scales	deterministic time evolution of probability distributions and mean values
	⇓ ⇒	short memory of fast variables has little impact on dynamics of slow variables at long times

Comments:

- In a classical Hamiltonian system the deterministic time evolution pertains to canonical coordinates and functions thereof.
- The time rate of change of any such variable depends on the instantaneous values of all canonical coordinates.
- On the contracted level of description we seek a way of describing an autonomous time evolution of a subset of variables.
- For that purpose the information contained in the instantaneous values of the variables that do not belong to the subset is transcribed into previous values of the variables that do belong to the subset.
- The autononmous time evolution of the variables belonging to the subset thus includes memory of its previous values.
- Slow variables contribute long memory and fast variables contribute short memory.
- If the subset contains all slow variables then any effects on its autonomous time evolution contributed by the remaining variables involve only short memory.
- Effects of short memory are more easily accounted for than effects of long memory.

Markov Process: General Attributes [nln54]

Specification of Markov process:

- $P(x, t_0)$ (initial probability distribution),
- $P(x_1, t_1|x_2, t_2)$ (conditional probability distribution).

The entire hierarchy of joint probability distributions (see [nln50]) can be generated from these two ingredients if the process is Markovian.

Two times $t_1 \geq t_2$:

$$P(x_1, t_1; x_2, t_2) = P(x_1, t_1|x_2, t_2)P(x_2, t_2).$$

Three times $t_1 \geq t_2 \geq t_3$:

$$\begin{aligned} P(x_1, t_1; x_2, t_2; x_3, t_3) &= P(x_1, t_1; x_2, t_2|x_3, t_3)P(x_3, t_3) \\ &= P(x_1, t_1|x_2, t_2; x_3, t_3)P(x_2, t_2|x_3, t_3)P(x_3, t_3) \\ &= P(x_1, t_1|x_2, t_2)P(x_2, t_2|x_3, t_3)P(x_3, t_3). \end{aligned}$$

Comments:

- The step from the first to the second line uses the previous equation in a reduced sample space (specified by one condition).
- The second condition in the middle line is redundant.

Integration over the variable x_2 at intermediate time t_2 yields

$$\underbrace{P(x_1, t_1; x_3, t_3)}_{P(x_1, t_1|x_3, t_3)P(x_3, t_3)} = P(x_3, t_3) \int dx_2 P(x_1, t_1|x_2, t_2)P(x_2, t_2|x_3, t_3).$$

Division by $P(x_3, t_3)$ then yields the *Chapman-Kolmogorov* equation:

$$P(x_1, t_1|x_3, t_3) = \int dx_2 P(x_1, t_1|x_2, t_2)P(x_2, t_2|x_3, t_3), \quad t_1 \geq t_2 \geq t_3.$$

The Chapman-Kolmogorov equation is a functional equation between conditional probability distributions with many different kinds of solutions.

Put differently ...

Any two non-negative and normalized functions $P(x,t)$ and $P(x_1,t_1|x_2,t_2)$ represent a unique Markov process if they satisfy the following two conditions:

- $P(x_1,t_1) = \int dx_2 P(x_1,t_1|x_2,t_2)P(x_2,t_2) \quad (t_1 \geq t_2)$,
- $P(x_1,t_1|x_3,t_3) = \int dx_2 P(x_1,t_1|x_2,t_2)P(x_2,t_2|x_3,t_3) \quad (t_1 \geq t_2 \geq t_3)$.

The first condition implies that $\lim_{\Delta t \to 0} P(x_1, t+\Delta t|x_2,t) = \delta(x_1 - x_2)$.

Homogeneous process: $P(x_1, t+\Delta t|x_0,t) \doteq P(x_1|x_0;\Delta t)$ independent of t.

The two conditions thus become

- $P(x_1, t+\Delta t) = \int dx_2 P(x_1|x_2;\Delta t)P(x_2,t)$,
- $P(x_1|x_3;\Delta t_{13}) = \int dx_2 P(x_1|x_2;\Delta t_{12})P(x_2|x_3;\Delta t_{23})$
  with $\Delta t_{13} = \Delta t_{12} + \Delta t_{23}$.

For initial condition $P(x,0) = \delta(x - x_0)$ we then have $P(x,t) = P(x|x_0;t)$.

Diffusion Process and Cauchy Process [nln55]

Here we portray two of the most common homogeneous Markov processes.

☐ Diffusion process: $P(x|x_0;\Delta t) = \dfrac{1}{\sqrt{4\pi D\Delta t}} \exp\left(-\dfrac{(x-x_0)^2}{4D\Delta t}\right)$.

☐ Cauchy process: $P(x|x_0;\Delta t) = \dfrac{1}{\pi}\dfrac{\Delta t}{(x-x_0)^2+(\Delta t)^2}$.

Both processes satisfy

- $\displaystyle\int_{-\infty}^{+\infty} dx P(x|x_0;\Delta t) = 1$ (normalization),
- $\displaystyle\lim_{\Delta t\to 0} P(x|x_0;\Delta t) = \delta(x-x_0)$ (consistency),
- $\displaystyle P(x_1|x_3;\Delta t_{13}) = \int dx_2 P(x_1|x_2;\Delta t_{12})P(x_2|x_3;\Delta t_{23})$ (C.-K. eq.) [nex26].

Q: Are the sample paths of the two processes continuous or discontinuous?

A: The sample paths are continuous in the diffusion process and discontinuous in the Cauchy process [nex97].

Lindeberg criterion for continuous sample paths:

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|x-x_0|>\epsilon} dx\, P(x|x_0;\Delta t) = 0$$

for any $\epsilon > 0$ and uniformly in x_0 and Δt.

Interpretation: the probability for the final position x to be finitely different from the initial position x_0 goes to zero faster than Δt as $\Delta t \to 0$.

Computer generated sample paths for both processes are shown in [nsl1].

Q: Are the sample paths of the two processe differentiable?

A: In both processes the sample paths are nowhere differentiable [nex99].

[nex26] Markovian nature of diffusion process and Cauchy process.

Demonstrate that the *diffusion* process and the *Cauchy* process are Markov processes by showing that the respective conditional probability densities

$$P(x|x_0;\Delta t) = \frac{1}{\sqrt{4\pi D\Delta t}}\,\exp\left(-\frac{(x-x_0)^2}{4D\Delta t}\right), \qquad P(x|x_0;\Delta t) = \frac{1}{\pi}\frac{\Delta t}{(x-x_0)^2+(\Delta t)^2}$$

satisfy the (integral) Chapman-Kolmogorov equation

$$P(x_1|x_3;\Delta t_{13}) = \int dx_2\, P(x_1|x_2;\Delta t_{12})P(x_2|x_3;\Delta t_{23}) \quad \text{with} \quad \Delta t_{13} = \Delta t_{12} + \Delta t_{23}.$$

What property of the characteristic function $\Phi(k,\Delta t) = \int d(x-x_0)e^{ik(x-x_0)}P(x|x_0;\Delta t)$ is instrumental in this context?

Solution:

Computer generated sample paths [nsl1]

Prescription:

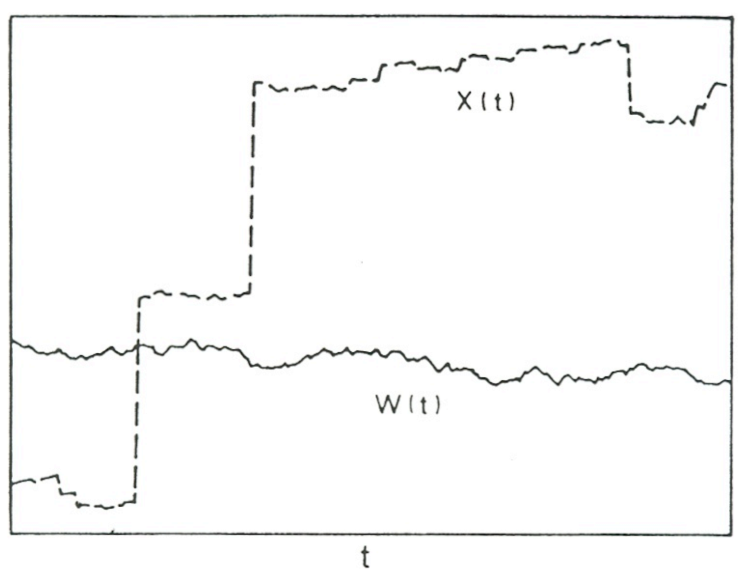
- Select Markov process: $P(x|x_0;\Delta t)$.
- Choose Δt sufficiently small (e.g. smaller than linewidth of graph).
- Produce sequence of (uniformly distributed) random numbers.
- Transform random numbers to fit $P(x|x_0;\Delta t)$ (see [nex80]).
- Use transformed random numbers as increments for sample path.

$W(t)$: Diffusion process (continuous) generated from

$$P(x|x_0;\Delta t) = \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left(-\frac{(x-x_0)^2}{4D\Delta t}\right).$$

$X(t)$: Cauchy process (discontinuous) generated from

$$P(x|x_0;\Delta t) = \frac{1}{\pi}\frac{\Delta t}{(x-x_0)^2 + (\Delta t)^2}.$$



[from Gardiner 1985]

[nex97] Lindeberg condition for diffusion and Cauchy processes

Show that the Lindeberg condition for continuity of sample paths,

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-x_0|>\epsilon} dx \, P(x|x_0; \Delta t) = 0,$$

is satisfied by the diffusion process but violated by th Cauchy process. They are specified, respectively, by the conditional probability distributions,

$$P(x|x_0; \Delta t) = \frac{1}{\sqrt{4\pi D \Delta t}} \exp\left(-\frac{(x-x_0)^2}{4D\Delta t}\right), \qquad P(x|x_0; \Delta t) = \frac{1}{\pi} \frac{\Delta t}{(x-x_0)^2 + (\Delta t)^2}.$$

The condition requires that the probability for the final position x to deviate a finite distance from the initial position x_0 vanishes faster than the time step Δt in the limit $\Delta t \to 0$.

Solution:

Differential Chapman-Kolmogorov Equation [nln56]

Focus on particular solutions of the (integral) Chapman-Kolmogorov equation that satisfy three conditions:

$$\text{(i)} \quad \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(x, t + \Delta t | x_0, t) = W(x|x_0; t) > 0,$$

$$\text{(ii)} \quad \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-x_0|<\epsilon} dx (x - x_0) P(x, t + \Delta t | x_0, t) = A(x_0, t) + \mathrm{O}(\epsilon),$$

$$\text{(iii)} \quad \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-x_0|<\epsilon} dx (x - x_0)^2 P(x, t + \Delta t | x_0, t) = B(x_0, t) + \mathrm{O}(\epsilon).$$

Comments:

- Integrals such as in (ii) and (iii) but with higher moments vanish,
- $W(x|x_0; t) > 0$ describes jumps,
- $A(x_0, t)$ describes drift,
- $B(x_0, t)$ describes diffusion.

Under assumptions including the ones stated above the following *differential Chapman-Kolmogorov equation* can be derived from its integral counterpart [see e.g. Gardiner 1985]:

$$\frac{\partial}{\partial t} P(x, t|x_0, t_0) = -\frac{\partial}{\partial x}\Big[A(x, t)P(x, t|x_0, t_0)\Big] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\Big[B(x, t)P(x, t|x_0, t_0)\Big]$$
$$+ \int dx' \Big[W(x|x'; t)P(x', t|x_0, t_0) - W(x'|x; t)P(x, t|x_0, t_0)\Big].$$

Initial condition: $P(x, t_0|x_0, t_0) = \delta(x - x_0)$.

Special cases:

- Drift equation: first term only. [nex29]
- Fokker-Planck equation: first and second terms only.
- Master equation: third term only.
- Diffusion process has $W = 0$, $A = 0$, $B \neq 0$. [nex27]
- Cauchy process has $W \neq 0$, $A = 0$, $B = 0$. [nex98]

Fokker-Planck Equation [nln57]

Extraction through systematic approximation of a (specific) Fokker-Planck equation from the (unspecific) Chapman-Kolmogorov equation for a homogeneous Markov process with continuous sample path.

Chapman-Kolmogorov equation: $P(x|x_0;t+\tau) = \int dx' P(x|x';\tau)P(x'|x_0;t)$.

Introduce a function $R(x)$ that is differentiable and vanishes at the boundaries of the range of x.

$$\int dx R(x) \frac{\partial}{\partial t} P(x|x_0;t) = \lim_{\tau\to 0} \frac{1}{\tau} \int dx R(x) \Big[P(x|x_0;t+\tau) - P(x|x_0;t) \Big] \quad (1)$$

$$= \lim_{\tau\to 0} \frac{1}{\tau} \int dx R(x) \left[\int dx' P(x|x';\tau)P(x'|x_0;t) - P(x|x_0;t) \right] \quad (2)$$

$$= \lim_{\tau\to 0} \frac{1}{\tau} \int dx\, P(x|x_0;t) \Big[\int dx' R(x')P(x'|x;\tau) - R(x) \Big] \quad (3)$$

$$= \lim_{\tau\to 0} \frac{1}{\tau} \int dx\, P(x|x_0;t) \int dx' P(x'|x;\tau) \times \left[(x'-x)R'(x) + \frac{1}{2}(x'-x)^2 R''(x) + \ldots \right] \quad (4)$$

$$= \lim_{\tau\to 0} \frac{1}{\tau} \int dx\, R(x) \left[-\frac{\partial}{\partial x} \int dx'(x'-x)P(x'|x;\tau)P(x|x_0;t) + \frac{1}{2}\frac{\partial^2}{\partial x^2} \int dx'(x'-x)^2 P(x'|x;\tau)P(x|x_0;t) + \ldots \right]. \quad (5)$$

(1) Construct partial time derivative;

(2) use Chapman-Kolmogorov equation;

(3) switch variables x, x';

(4) expand function $R(x)$ at position x;

(5) integrate by parts.

Since the above equation must hold for any $R(x)$ with the attributes mentioned the Fokker-Planck equation follows.

$$\frac{\partial}{\partial t} P(x|x_0;t) = -\frac{\partial}{\partial x} A(x)P(x|x_0;t) + \frac{1}{2}\frac{\partial^2}{\partial x^2} B(x)P(x|x_0;t)$$

with drift and diffusion coefficients

$$A(x) = \lim_{\tau\to 0} \frac{1}{\tau} \int dx'(x'-x)P(x'|x;\tau),$$

$$B(x) = \lim_{\tau\to 0} \frac{1}{\tau} \int dx'(x'-x)^2 P(x'|x;\tau).$$

[nex29] Drift equation.

The Fokker-Planck equation with no diffusive term reads

$$\frac{\partial}{\partial t} P(x,t|x_0,0) = -\frac{\partial}{\partial x}\,[A(x,t)P(x,t|x_0,0)],$$

Show that this drift equation has has a solution of the form

$$P(x,t|x_0,0) = \delta\big(x - x_S(t)\big),$$

where $x_S(t)$ is the solution of the (deterministic) equation of motion $dx/dt = A(x,t)$ with initial condition $x_S(0) = x_0$.

Comment: In the $6N$-dimensional phase space of a classical system of N interacting particles, the drift equation for a phase point $\mathbf{x}$ is the Liouville equation. There exists a solution of the form $\delta\big(\mathbf{x} - \mathbf{x}_S(t)\big)$, representing the motion of a phase point through phase space. The function $\mathbf{x}_S(t)$ is the solution of the canonical equations, which have the form $d\mathbf{x}/dt = \mathbf{A}(\mathbf{x},t)$. The function $\mathbf{A}(\mathbf{x},t)$ is constructed from the Poisson bracket of $\mathbf{x}$ with the system Hamiltonian.

Solution:

[nex28] Master equation for a continuous random variable.

The master equation,

$$\frac{\partial}{\partial t} P(x|x_0;t) = \int dx' [W(x|x')P(x'|x_0;t) - W(x'|x)P(x|x_0;t)],$$

describes the time evolution of probability distributions for pure jump processes. The rate at which $P(x|x_0;t)$ evolves in time is governed by two contributions: a positive contribution from jumps $x' \to x$ taking place at the rate $W(x|x')$ and a negative contribution from jumps $x \to x'$ taking place at the rate $W(x'|x)$. The goal of this exercise is to derive the master equation from the ansatz

$$P(x|x';\Delta t) = \Delta t\, W(x|x') + \delta(x-x')[1 - \Delta t \int dx'' W(x''|x')],$$

for the conditional probability distribution assumed to hold for infinitesimal time intervals Δt. In this ansatz, the first term represents the probability density for transitions $x' \to x \neq x'$ during Δt and the second term the probability density for no transitions occurring within Δt.

Hint: Start from the (integral) Chapman-Kolmogrov equation for $P(x|x_0;t)$ and construct the partial time derivative via $\lim_{\Delta t \to 0}[P(x|x_0;t+\Delta t) - P(x|x_0;t)]/\Delta t$. Then substitute the ansatz.

Solution:

[nex99] Non-differentiability of sample paths

The non-differentiability of sample paths of stochastic processes can be investigated by calculating the probability that for any value of x the slope $\lim_{\Delta t\to 0}|\Delta x/\Delta t|$ is greater than any chosen value $\kappa > 0$. Calculate

$$\lim_{\Delta t\to 0}\mathcal{P}[|\Delta x/\Delta t| > \kappa] = \int_{|x-x_0|>\kappa\Delta t} dx\, P(x|x_0;\Delta t)$$

for the diffusion process and the Cauchy process,

$$P(x|x_0;\Delta t) = \frac{1}{\sqrt{4\pi D\Delta t}}\exp\left(-\frac{(x-x_0)^2}{4D\Delta t}\right), \qquad P(x|x_0;\Delta t) = \frac{1}{\pi}\frac{\Delta t}{(x-x_0)^2+(\Delta t)^2}.$$

For comparison, calculate the same probability for the deterministic drift process,

$$P(x|x_0;\Delta t) = \delta(x - x_0 - v\Delta t).$$

Draw your own conclusions from the results. State a necessary criterion for a sample path to be differentiable at a given value of x. Which process satisfies your criterion?

Solution:

Predominantly Small Jumps [nln58]

Jump processes are most commonly described by a master equation,

$$\frac{\partial}{\partial t} P(x,t|x_0) = \int dx' [W(x|x')P(x',t|x_0) - W(x'|x)P(x,t|x_0)].$$

If the transition rates favor small jumps such that their expansion in powers of jump size captures the essence of the process at hand we can extract a Fokker-Planck equation from the master equation.

Express transition rates as functions of jump size $\xi \doteq x' - x$:

$$W(x'|x) = \bar{W}(x;\xi), \quad W(x|x') = \bar{W}(x';-\xi).$$

Rewrite master equation:

$$\frac{\partial}{\partial t}P(x,t|x_0) = \int d\xi \Big[\bar{W}(x+\xi;-\xi)P(x+\xi,t|x_0) - \bar{W}(x;\xi)P(x,t|x_0)\Big].$$

Expand first term to second order:

$$\bar{W}(x;-\xi)P(x,t|x_0) + \xi\frac{\partial}{\partial x}\big[\bar{W}(x;-\xi)P(x,t|x_0)\big] + \frac{1}{2}\xi^2\frac{\partial^2}{\partial x^2}\big[\bar{W}(x;-\xi)P(x,t|x_0)\big].$$

Introduce jump moments:

$$\alpha_m(x) \doteq \int d\xi\, \xi^m \bar{W}(x;\xi) = \int dx'(x'-x)^m W(x'|x).$$

Substitution of expansion into master equation yields Fokker-Planck equation:

$$\frac{\partial}{\partial t}P(x,t|x_0) = -\frac{\partial}{\partial x}\big[\alpha_1(x)P(x,t|x_0)\big] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\big[\alpha_2(x)P(x,t|x_0)\big].$$

Comments:

- Convergent jump moments necessitate predominance of small jumps.
- The jump moments $\alpha_1(x)$ and $\alpha_2(x)$ only capture partial information contained in the transition rates $W(x|x')$, namely information associated with effective drift and effective diffusion.

Time Evolution of Mean and Variance [nln59]

Consider a stochastic process specified by the master equation

$$\frac{\partial}{\partial t}P(x,t|x_0) = \int dx'[W(x|x')P(x',t|x_0) - W(x'|x)P(x,t|x_0)].$$

Jump moments are extracted from the transition rates via

$$\alpha_m(x) \doteq \int dx'(x'-x)^m W(x'|x)$$

and assumed to be convergent at least for $m = 1,2$ (see [nln58]).

Evaluate $\int dx x$[m.eq.] and $\int dx x^2$[m.eq.] to express the rate at which the first and second moments of x vary in time as follows:

$$\frac{d}{dt}\langle x\rangle = \int dx \int dx'(x'-x)W(x'|x)P(x,t|x_0),$$

$$\frac{d}{dt}\langle x^2\rangle = \int dx \int dx'(x'^2-x^2)W(x'|x)P(x,t|x_0).$$

Use $x'^2 - x^2 = (x'-x)^2 + 2x(x'-x)$ and the definition of jump moments to derive the following equations of motion:

$$\frac{d}{dt}\langle x\rangle = \langle \alpha_1(x)\rangle,$$

$$\frac{d}{dt}\langle x^2\rangle = \langle \alpha_2(x)\rangle + 2\langle x\alpha_1(x)\rangle.$$

Comments:

- If the jump moments are known and expandable in powers of x the expectation values on the right-hand sides become functions of $\langle x\rangle, \langle x^2\rangle, \ldots$
- In general, this leads to an infinite hierarchy of equations of motion for all moments $\langle x^m\rangle, m = 1,2,\ldots$.
- In special cases, the ODEs for $\langle x\rangle$ and $\langle x^2\rangle$ form a closed set. Then they can be solved with no further approximations.
- The same equations of motions hold if the first two jump moments are replaced by the drift and diffusion coefficents of a Fokker-Planck equation, $A(x)$ and $B(x)$, respectively (see nln58]).
- Solvable cases are worked out in [nex30], [nex32].

[nex32] Jump moments of discrete variables

Consider the master equation

$$\frac{d}{dt}P(n,t) = \sum_m [W(n|m)P(m,t) - W(m|n)P(n,t)]$$

of an integer random variable n for two stochastic processes:
(a) Random walk: $W(n|m) = \sigma\delta_{n+1,m} + \sigma\delta_{n-1,m}$.
(b) Poisson process: $W(n|m) = \lambda\delta_{n-1,m}$.
Calculate the jump moments $\alpha_l(m) = \sum_n (n-m)^l W(n|m)$ for $l = 1, 2$.
Then calculate the time evolution of the mean value $\langle n \rangle$ and the variance $\langle\langle n^2 \rangle\rangle$, consistent with the initial condition $P(n,0) = \delta_{n,0}$. Rather than first calculating $P(n,t)$, solve the equations of motion for the expectation values: $d\langle n \rangle/dt = \langle \alpha_1(n) \rangle$, $d\langle n^2 \rangle/dt = \langle \alpha_2(n) \rangle + 2\langle n\alpha_1(n) \rangle$.

Solution:

[nex30] Equations of motion for mean value and variance.

Consider the Fokker-Planck equation for a stochastic process,

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

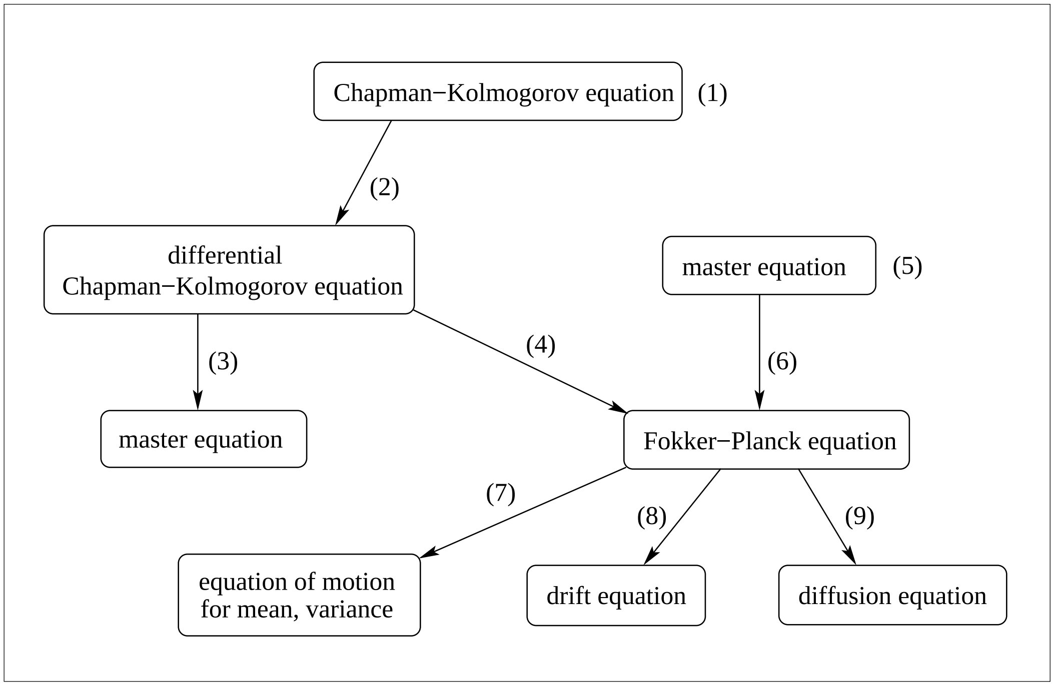
where x_0 is the initial value of all sample paths, implying $P(x,0|x_0) = \delta(x-x_0)$. Use the equations of motion,

$$\frac{d}{dt}\,\langle x\rangle = \langle A(x)\rangle, \qquad \frac{d}{dt}\,\langle x^2\rangle = \langle B(x)\rangle + 2\langle xA(x)\rangle,$$

to calculate the time-dependence of the mean value $\langle x\rangle$ and the variance $\langle\langle x^2\rangle\rangle$ for two processes with initial conditions as dictated by $P(x,0|x_0) = \delta(x-x_0)$:
(i) Uniform drift and diffusion process: $A(x) = v$, $B(x) = 2D$.
(ii) Ornstein-Uhlenbeck process: $A(x) = -\kappa x$, $B(x) = \gamma$.

Solution:

Markov processes: map of specifications [nln16]



(1) Chapman-Kolmogorov equation imposes restrictions on permissible functions $P(x,t|x_0)$ but does not suggest a classification of processes.

(2) Particular solutions that are specified by
   - $A(x,t)$ describing drift,
   - $B(x,t)$ describing diffusion,
   - $W(x|x';t)$ describing jumps.

(3) Jump processes exclusively.

(4) Processes with continuous sample paths, satisfying Lindeberg criterion (drift and diffusion, no jumps).

(5) Master equation with any $W(x|x';t)$ specifies a Markov process. Natural starting point for processes with discrete stochastic variables.

(6) Transition rates $W(x|x';t)$ of master equation approximated by two jump moments provided they exist. Approximation captures drift and diffusion parts (on some scale).

(7) Drift and diffusion determine mean $\langle\langle x\rangle\rangle$ and variance $\langle\langle x^2\rangle\rangle$ via equations of motion for jump moments.

(8) Deterministic process have no diffusive part: $B(x,t) = 0$.

(9) Purely diffusive processes have no drift: $A(x,t) = 0$.

[nex85] Detailed balance condition and thermal equilibrium

Consider a statistical mechanical system specified by a Hamiltonian $H(\mathbf{x})$. Here the *random field* $\mathbf{x} = (x_1, x_2, \ldots)$ specifies the microstate. At thermal equilibrium, the macrostate is specified by the probability distribution $\rho(\mathbf{x}) = Z^{-1}\exp[-\beta H(\mathbf{x})]$. Now consider a Markov process specified by the master equation

$$\frac{\partial}{\partial t}P(\mathbf{x},t) = \sum_{\mathbf{x}'}\left[W(\mathbf{x}|\mathbf{x}')P(\mathbf{x}',t) - W(\mathbf{x}'|\mathbf{x})P(\mathbf{x},t)\right].$$

Show that the equilibrium distribution $\rho(\mathbf{x})$ satisfies the detailed balance condition $W(\mathbf{x}'|\mathbf{x})\rho(\mathbf{x}) = W(\mathbf{x}|\mathbf{x}')\rho(\mathbf{x}')$ for the *Metropolis algorithm* and the *heat bath algorithm*, which are specified, respectively by the transition rates $[\Delta H \equiv H(\mathbf{x}') - H(\mathbf{x})]$:

$$W(\mathbf{x}'|\mathbf{x})dt = \begin{cases} e^{-\beta\Delta H} & \text{if } \Delta H \geq 0 \\ 1 & \text{if } \Delta H \leq 0 \end{cases}, \qquad W(\mathbf{x}'|\mathbf{x})dt = \frac{e^{-\beta\Delta H}}{1+e^{-\beta\Delta H}}.$$

Solution:

Markov Chains [nln61]

Transitions between values of a discrete stochastic variable taking place at discrete times:

$$X = \{x_1, \ldots, x_N\}; \qquad t = s\tau, \quad s = 0, 1, 2, \ldots$$

Notation adapted to accommodate linear algebra:

$$P(x_n, t) \to P(n, s), \quad P(x_n, t_0 + s\tau | x_m, t_0) \to P(n|m; s).$$

Time evolution of initial probability distribution:

$$P(n, s) = \sum_m P(n|m; s) P(m, 0).$$

Nested Chapman-Kolmogorov equations:

$$\begin{aligned} P(n|m; s) &= \sum_i P(n|i; 1) P(i|m; s-1) \\ &= \sum_{ij} P(n|i; 1) P(i|j; 1) P(j|m; s-2) \\ &= \sum_{ijk} P(n|i; 1) P(i|j; 1) P(j|k; 1) P(k|m; s-3) = \ldots \end{aligned}$$

Matrix representation:

Transition matrix: $\mathbf{W}$ with elements $W_{mn} = P(n|m; 1)$.

Probability vector: $\vec{P}(s) = \Big(P(1, s), \ldots, P(N, s)\Big)$.

Time evolution via matrix multiplication: $\vec{P}(s) = \vec{P}(0) \cdot \mathbf{W}^s$.

General attributes of transition matrix:

- All elements represent probabilities: $W_{mn} \geq 0$;
  W_{mm}: system stays in state m;
  W_{mn} with $m \neq n$: system undergoes a transition from m to n.
- Normalization of probabilities: $\sum_n W_{mn} = 1$
- A transition $m \to n$ and its inverse $n \to m$ may occur at different rates. Hence $\mathbf{W}$ is, in general, not symmetric.

Regularity:

A transition matrix $\mathbf{W}$ is called *regular* if all elements of the matrix product $\mathbf{W}^s$ are nonzero (i.e. positive) for some exponent s.

Regularity guarantees that repeated multiplication leads to convergence:

$$\lim_{s\to\infty} \mathbf{W}^s = \mathbf{M} = \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_N \\ \pi_1 & \pi_2 & \cdots & \pi_N \\ \vdots & \vdots & & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_N \end{pmatrix}$$

Further multiplications have no effect:

$$\mathbf{W} \cdot \mathbf{M} = \begin{pmatrix} W_{11} & \cdots & W_{1N} \\ \vdots & & \vdots \\ W_{N1} & \cdots & W_{NN} \end{pmatrix} \cdot \begin{pmatrix} \pi_1 & \cdots & \pi_N \\ \vdots & & \vdots \\ \pi_1 & \cdots & \pi_N \end{pmatrix} = \mathbf{M}.$$

The asymptotic distribution is stationary.
The stationary distribution does not depend on initial distribution:

$$\lim_{s\to\infty} \vec{P}(s) = \vec{P}(0) \cdot \mathbf{M} = \vec{\pi} = \Big(\pi_1, \pi_2, \ldots, \pi_N\Big).$$

All elements of the stationary distribution are nonzero.

The computation of the stationary distribution $\vec{\pi}$ via repeated multiplication of the transition matrix with itself works well for regular matrices.

More generally, transition matrices may have stationary solutions that depend on the initial distribution or stationary solutions that are not asymptotic solutions of any kind.

Eigenvalue problem:

The eigenvalues $\Lambda_1, \ldots, \Lambda_N$ of $\mathbf{W}$ are the solutions of the secular equation:

$$\det\big(\mathbf{W} - \Lambda\mathbf{E}\big) = 0, \quad E_{ij} = \delta_{ij}.$$

For an asymmetric $\mathbf{W}$ not all eigenvalues Λ_n are real. We must distinguish between left eigenvectors $\vec{X}_n$ and right eigenvectors $\vec{Y}_n$:

$$\vec{X}_n \cdot \mathbf{W} = \Lambda_n \vec{X}_n, \quad n = 1, \ldots, N \quad \text{with} \quad \vec{X}_n \doteq \big(X_{n1}, \ldots, X_{nN}\big)$$

$$\mathbf{W} \cdot \vec{Y}_n = \Lambda_n \vec{Y}_n, \quad n = 1, \ldots, N \quad \text{with} \quad \vec{Y}_n = \begin{pmatrix} Y_{1n} \\ \vdots \\ Y_{Nn} \end{pmatrix}.$$

The two eigenvector matrices are orthonormal to one another:

$$\mathbf{X} \cdot \mathbf{Y} = \mathbf{E}, \quad \text{where} \quad \mathbf{X} \doteq \begin{pmatrix} \vec{X}_1 \\ \vdots \\ \vec{X}_N \end{pmatrix}, \quad \mathbf{Y} \doteq \big(\vec{Y}_1, \ldots, \vec{Y}_N\big).$$

All eigenvalues Λ_n of the transition matrix $\mathbf{W}$ satisfy the condition $|\Lambda_n| \leq 1$.

There always exists at least one eigenvalue $\Lambda_n = 1$.

The right eigenvector for $\Lambda_n = 1$ is $\vec{Y}_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

The left eigenvector for $\Lambda_n = 1$ is a stationary distribution $\vec{X}_n = \big(\pi_1, \ldots, \pi_N\big)$.

If $\mathbf{W}$ is regular then the eigenvalue $\Lambda_n = 1$ is unique and its left eigenvector is the asymptotic distribution $\vec{X}_n = \vec{\pi}$, independent of the inital condition.

Ergodicity:

In an *ergodic* transition matrix $\mathbf{W}$ any two states are connected, directly or indirectly, by allowed transitions. Regularity implies ergodicity but not vice versa.

A block-diagonal transition matrix,

$$\mathbf{W} = \begin{pmatrix} W_{1,1} & \cdots & W_{1,n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ W_{n,1} & \cdots & W_{n,n} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & W_{n+1,n+1} & \cdots & W_{n+1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & W_{N,n+1} & \cdots & W_{N,N} \end{pmatrix}$$

implies non-ergodicity because inter-block transitions are prohibited.

Absorbing states:

If there exists a state n that allows only transitions into it but not out of it then row n of the transition matrix has diagonal element $W_{nn} = 1$ and off-diagonal elements $W_{nn'} = 0$ ($n' \neq n$).

For an ergodic system we then have

$$\lim_{s\to\infty} \vec{P}(0) \cdot \mathbf{W}^s = \vec{\pi} = (0, \ldots, 0, 1, 0, \ldots, 0),$$

with the 1 at position n.

Detailed balance:

The detailed balance condition postulates the existence of a stationary distribution $\vec{\pi}$ satisfying the relations

$$W_{mn}\pi_m = W_{nm}\pi_n, \quad n, m = 1, \ldots, N.$$

Detailed balance requires that $W_{mn} = 0$ if $W_{nm} = 0$. Microscopic (quantum or classical) dynamics guarantees that this requirement is fulfilled.

The detailed balance condition, if indeed satisfied, can be used to determine the stationary distribution.

Applications:

- ▷ House of the mouse: two-way doors only [nex102]
- ▷ House of the mouse: some one-way doors [nex103]
- ▷ House of the mouse: one-way doors only [nex104]
- ▷ House of the mouse: mouse with inertia [nex105]
- ▷ House of the mouse: mouse with memory [nex43]
- ▷ Mixing marbles red and white [nex42]
- ▷ Random traffic around city block [nex86]
- ▷ Modeling a Markov chain [nex87]

Master Equation with detailed balance [nln12]

Master equation with time-independent transition rates $W(n|m) = W_{mn}$:

$$\frac{\partial}{\partial t} P(n,t) = \sum_m \left[W_{mn} P(m,t) - W_{nm} P(n,t) \right] = \sum_m L_{mn} P(m,t),$$

where $L_{mn} = W_{mn} - \delta_{mn} \sum_{n'} W_{nn'} = W_{mn} - \delta_{mn}$.

This set of linear, ordinary differential equations can be transformed into an eigenvalue problem with the ansatz $P(m,t) = \varphi_m e^{-\lambda t}$:

Left eigenvector problem: $\displaystyle\sum_m L_{mn} \varphi_m^{(\alpha)} = -\lambda^{(\alpha)} \varphi_n^{(\alpha)}, \quad \alpha = 1, 2, \ldots$

Right eigenvector problem: $\displaystyle\sum_n L_{mn} \chi_n^{(\alpha)} = -\lambda^{(\alpha)} \chi_m^{(\alpha)}, \quad \alpha = 1, 2, \ldots$

Biorthonormality: $\displaystyle\vec{\varphi}^{(\alpha)} \cdot \vec{\chi}^{(\beta)} = \sum_n \varphi_n^{(\alpha)} \chi_n^{(\beta)} = \delta_{\alpha\beta}$.

A stationary solution $P(n)$ requires the existence of a solution of the eigenvalue problem with $\lambda = 0$. The stability of $P(n)$ requires that all other eigenvalues λ have positive real parts.

Detailed balance condition: $W_{mn} P(m) = W_{nm} P(n)$.

Symmetric matrix: $\displaystyle S_{mn} \doteq L_{mn} \sqrt{\frac{P(m)}{P(n)}}$.

Symmetrized eigenvalue problem:

$$\bar{\varphi}_n^{(\alpha)} \doteq \frac{1}{\sqrt{P(n)}} \varphi_n^{(\alpha)} \quad \Rightarrow \quad \sum_m S_{mn} \bar{\varphi}_m^{(\alpha)} = -\lambda^{(\alpha)} \bar{\varphi}_n^{(\alpha)}.$$

$$\bar{\chi}_n^{(\alpha)} \doteq \sqrt{P(n)}\, \chi_n^{(\alpha)} \quad \Rightarrow \quad \sum_n S_{mn} \bar{\chi}_n^{(\alpha)} = -\lambda^{(\alpha)} \bar{\chi}_m^{(\alpha)}.$$

Given that $\bar{\varphi}_n^{(\alpha)} = \bar{\chi}_n^{(\alpha)}$ it follows that $\varphi_n^{(\alpha)} = P(n) \chi_n^{(\alpha)}$.

Given that $\sum_n L_{mn} = 0$ it follows that the right eigenvector of L_{mn} with eigenvalue $\lambda = 0$ has components $\chi_n = 1$. The corresponding left eigenvector then has components $\varphi_n = P(n)$.

The symmetric matrix $\mathbf{S}$ has only real, non-negative eigenvalues. Hence $\lambda = 0$ is the smallest eigenvalue. Variational methods are applicable.

[nex39] Regression theorem for autocorrelation functions.

The regression theorem for autocorrelation functions of a Markov process reads

$$\langle X(t)X(t_0)\rangle = \int dx \int dx' x x_0 P(x,t;x_0,t_0) = \int dx_0 \langle X(t)|[x_0,t_0]\rangle x_0 P(x_0,t),$$

where $\langle X(t)|[x_0,t_0]\rangle \equiv \int dx\, x P(x,t|x_0,t_0)$ is the definition of a conditional average.

(a) Show that if $\lim_{t_0\to-\infty} P(x,t|x_0,t_0) = P_S(x)$ independent of t, x_0, then the autocorrelation function in a stationary process is

$$\langle X(t)X(t')\rangle_S = \lim_{t_0\to-\infty} \langle X(t)X(t')|[x_0,t_0]\rangle = \int dx \int dx' x x' P(x,t|x',t')P_S(x').$$

(b) Apply the regression theorem to calculate $\langle X(t)X(t')\rangle_S$ for the Ornstein-Uhlenbeck process at stationarity.

Solution:

Birth-death processes [nln18]

Specification:

- Birth rates: typically proportional to population present.
- Death rates: typically proportional to population present, may be enhanced due to self-inflicted stress.
- Interaction rates: typically proportional to products of populations, with positive sign if impact is favorable and negative sign if impact is unfavorable.

Models for population dynamics:

- particles diffusing through walls,
- particles undergoing radioactive decay,
- molecules undergoing chemical reactions,
- organisms multiplying and dying,
- host-parasite interaction,
- predator-prey interaction,
- animals subject to environmental stress

Levels of description:

- Deterministic time evolution.
  Description via differential equations.
  Contingency encoded in initial conditions.
- Probabilistic time evolution without memory.
  Description via master equation.
  Contingency encoded in initial conditions and in time evolution.
- Probabilistic time evolution with memory.
  Incorporation of learning, heredity, adaptation.
  Contingency encoded in initial conditions, in time evolution, and in rules that govern time evolution.

The future is open to a higher degree in each successive level of description.

Birth and death of single species [nln19]

Class of processes described by a master equation for some discrete variable n with nonzero transition rates $W(m|n)$ limited to $m = n+1$ and $m = n-1$:

$$\frac{d}{dt}P(n,t) = \sum_m \Big[W(n|m)P(m,t) - W(m|n)P(n,t)\Big],$$

$$W(m|n) = \underbrace{T_+(n)\delta_{m,n+1}}_{\text{birth rate}} + \underbrace{T_-(n)\delta_{m,n-1}}_{\text{death rate}}.$$

The master equation is a difference-differential equation. If $T_\pm(n)$ are polynomials, the master equation can be converted into a linear PDE for the generating function $G(z,t) \doteq \sum_n z^n P(n,t)$:

$$\frac{\partial}{\partial t}G(z,t) = \sum_{l=0}^{L} A_l(z)\frac{\partial^l}{\partial z^l}G(z,t),$$

where L is the highest polynomial order in $T_\pm(n)$.

The notion of nonlinear birth/death rates pertains to quadratic or higher-order terms in $T_\pm(n)$. The PDE for $G(z,t)$ and the master equation for $P(n,t)$ remain linear. The relative ease of solving *linear* birth-death processes is associated with the relative ease of solving *first-order* linear PDEs.

In the context of a deterministic description of the time evolution, nonlinear birth/death rates translate into nonlinear differential equations.

Not all choices of transition rates $T_\pm(n)$ permit a stationary solution,

$$\lim_{t\to\infty} P(n,t) = P_s(n).$$

- Runaway populations can be held in check by death rates that are of higher polynomial order than the birth rates.
- Extinction of populations can be held in check by allowing births out of zero population.

Birth-death master equation: stationary state [nln17]

Master equation: $\displaystyle \frac{d}{dt}P(n,t) = \sum_m \Big[W(n|m)P(m,t) - W(m|n)P(n,t)\Big].$

Transition rates: $W(m|n) = \underbrace{T_+(n)\delta_{m,n+1}}_{\text{birth rate}} + \underbrace{T_-(n)\delta_{m,n-1}}_{\text{death rate}}.$

$$\Rightarrow \frac{d}{dt}P(n,t) = T_+(n-1)P(n-1,t) + T_-(n+1)P(n+1,t) - \big[T_+(n) + T_-(n)\big]P(n,t).$$

Stationary state: $P(n,\infty) = P_s(n)$.

Detailed-balance condition: $T_-(n)P_s(n) = T_+(n-1)P_s(n-1), \quad n = 0, 1, 2, \ldots$

Recurrence relation: $\displaystyle P_s(n) = \frac{T_+(n-1)}{T_-(n)}\, P_s(n-1).$

Prerequisites:

- $T_-(0) = 0$ (no further deaths at zero population),
- $T_+(0) > 0$ (spontaneous birth from nothing must be permitted if death of last individual is permitted).

Solution: $\displaystyle P_s(n) = P_s(0) \prod_{m=1}^{n} \frac{T_+(m-1)}{T_-(m)}.$

Probability of zero population, $P_s(0)$, determined by normalization condition:

$$\sum_{n=0}^{\infty} P_s(n) = 1.$$

Condition for extreme values (e.g. peak position) in $P_s(n)$:

$$T_+(n-1) = T_-(n).$$