# Lecture 1: Pragmatic Introduction to Stochastic Differential Equations

Simo Särkkä

Aalto University, Finland

October 28, 2014

# Contents

1. Introduction
2. Stochastic processes in physics and engineering
3. Heuristic solutions of linear SDEs
4. Fourier analysis of LTI SDEs
5. Heuristic solutions of non-linear SDEs
6. Summary and demonstration

# What is a stochastic differential equation (SDE)?

- At first, we have an ordinary differential equation (ODE):

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, t).$$

- Then we add white noise to the right hand side:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, t) + \mathbf{w}(t).$$

- Generalize a bit by adding a multiplier matrix on the right:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, t) + \mathbf{L}(\mathbf{x}, t)\,\mathbf{w}(t).$$

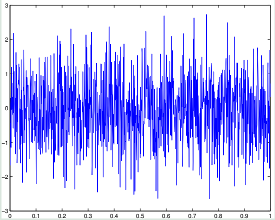
- Now we have a stochastic differential equation (SDE).
- $\mathbf{f}(\mathbf{x}, t)$ is the drift function and $\mathbf{L}(\mathbf{x}, t)$ is the dispersion matrix.

# White noise

## White noise

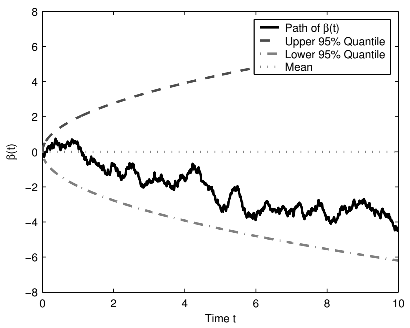
1. $\mathbf{w}(t_1)$ and $\mathbf{w}(t_2)$ are independent if $t_1 \neq t_2$.
2. $t \mapsto \mathbf{w}(t)$ is a Gaussian process with the mean and covariance:

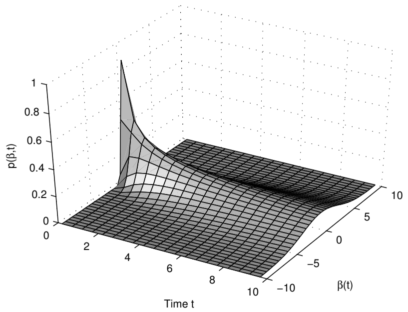
$$\mathrm{E}[\mathbf{w}(t)] = \mathbf{0}$$
$$\mathrm{E}[\mathbf{w}(t)\,\mathbf{w}^\mathsf{T}(s)] = \delta(t - s)\,\mathbf{Q}.$$




- $\mathbf{Q}$ is the spectral density of the process.
- The sample path $t \mapsto \mathbf{w}(t)$ is discontinuous almost everywhere.
- White noise is unbounded and it takes arbitrarily large positive and negative values at any finite interval.

# What does a solution of SDE look like?



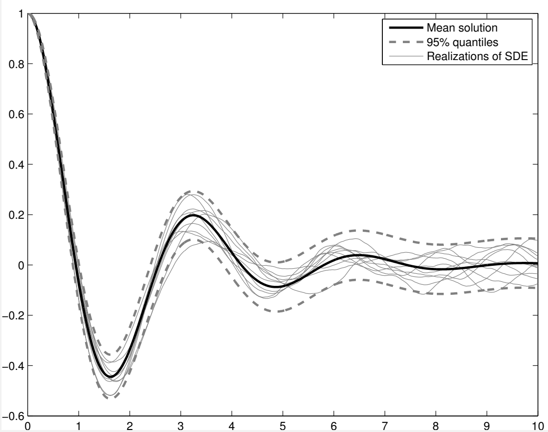




- *Left*: Path of a Brownian motion which is solution to stochastic differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = w(t)$$

- *Right*: Evolution of probability density of Brownian motion.

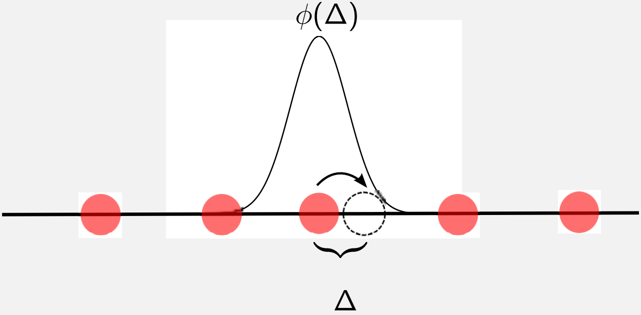
# What does a solution of SDE look like? (cont.)



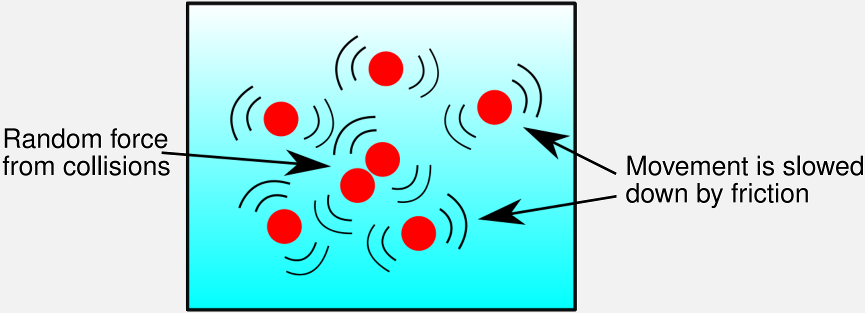

Paths of stochastic spring model

$$\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} + \gamma \frac{\mathrm{d}x(t)}{\mathrm{d}t} + \nu^2 x(t) = w(t).$$

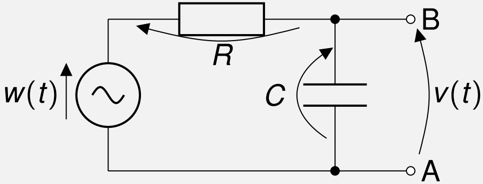
# Einstein’s construction of Brownian motion



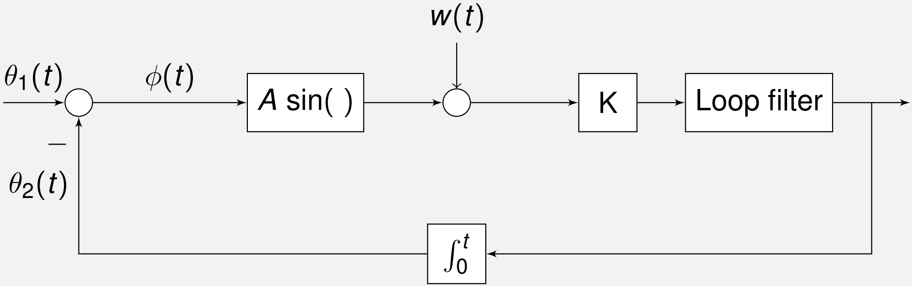
# Langevin's construction of Brownian motion



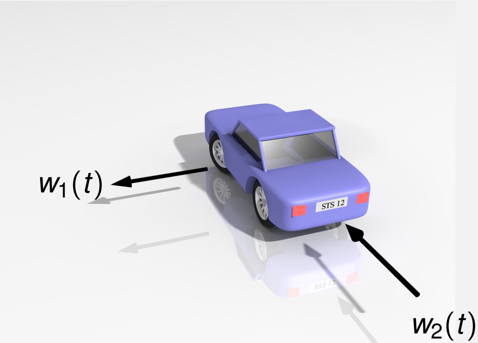
# Noisy RC-circuit



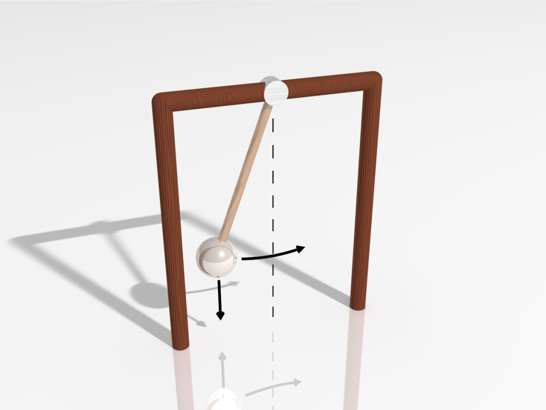
# Noisy Phase Locked Loop (PLL)



# Car model for navigation



# Noisy pendulum model



# Solutions of LTI SDEs

- Linear time-invariant stochastic differential equation (LTI SDE):

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t), \qquad \mathbf{x}(t_0) \sim \mathrm{N}(\mathbf{m}_0, \mathbf{P}_0).$$

- We can now take a "leap of faith" and solve this as if it was a deterministic ODE:
  1. Move $\mathbf{F}\,\mathbf{x}(t)$ to left and multiply by integrating factor $\exp(-\mathbf{F}\,t)$:

$$\exp(-\mathbf{F}\,t)\,\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} - \exp(-\mathbf{F}\,t)\,\mathbf{F}\,\mathbf{x}(t) = \exp(-\mathbf{F}\,t)\,\mathbf{L}\,\mathbf{w}(t).$$

  2. Rewrite this as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\exp(-\mathbf{F}\,t)\,\mathbf{x}(t)\right] = \exp(-\mathbf{F}\,t)\,\mathbf{L}\,\mathbf{w}(t).$$

  3. Integrate from $t_0$ to $t$:

$$\exp(-\mathbf{F}\,t)\,\mathbf{x}(t) - \exp(-\mathbf{F}\,t_0)\,\mathbf{x}(t_0) = \int_{t_0}^{t} \exp(-\mathbf{F}\,\tau)\,\mathbf{L}\,\mathbf{w}(\tau)\,\mathrm{d}\tau.$$

# Solutions of LTI SDEs (cont.)

- Rearranging then gives the solution:

$$\mathbf{x}(t) = \exp(\mathbf{F}\,(t - t_0))\,\mathbf{x}(t_0) + \int_{t_0}^{t} \exp(\mathbf{F}\,(t - \tau))\,\mathbf{L}\,\mathbf{w}(\tau)\;\mathrm{d}\tau.$$

- We have assumed that $\mathbf{w}(t)$ is an ordinary function, which it is not.
- Here we are lucky, because for linear SDEs we get the right solution, but generally not.
- The source of the problem is the integral of a non-integrable function on the right hand side.

# Mean and covariance of LTI SDEs

- The mean can be computed by taking expectations:

$$\mathrm{E}\,[\mathbf{x}(t)] = \mathrm{E}\,[\exp(\mathbf{F}\,(t - t_0))\,\mathbf{x}(t_0)] + \mathrm{E}\left[\int_{t_0}^{t} \exp(\mathbf{F}\,(t - \tau))\,\mathbf{L}\,\mathbf{w}(\tau)\;\mathrm{d}\tau\right]$$

- Recalling that $\mathrm{E}[\mathbf{x}(t_0)] = \mathbf{m}_0$ and $\mathrm{E}[\mathbf{w}(t)] = 0$ then gives the mean

$$\mathbf{m}(t) = \exp(\mathbf{F}\,(t - t_0))\,\mathbf{m}_0.$$

- We also get the following covariance (see the exercises...):

$$\begin{aligned}
\mathbf{P}(t) &= \mathrm{E}\left[(\mathbf{x}(t) - \mathbf{m}(t))\,(\mathbf{x}(t) - \mathbf{m})^{\mathsf{T}}\right] \\
&= \exp\left(\mathbf{F}\,t\right)\,\mathbf{P}_0\,\exp\left(\mathbf{F}\,t\right)^{\mathsf{T}} \\
&\quad + \int_0^t \exp\left(\mathbf{F}\,(t - \tau)\right)\,\mathbf{L}\,\mathbf{Q}\,\mathbf{L}^{\mathsf{T}}\,\exp\left(\mathbf{F}\,(t - \tau)\right)^{\mathsf{T}}\,\mathrm{d}\tau.
\end{aligned}$$

# Mean and covariance of LTI SDEs (cont.)

- By differentiating the mean and covariance expression we can derive the following differential equations for the mean and covariance:

$$\frac{\mathrm{d}\mathbf{m}(t)}{\mathrm{d}t} = \mathbf{F}\,\mathbf{m}(t)$$
$$\frac{\mathrm{d}\mathbf{P}(t)}{\mathrm{d}t} = \mathbf{F}\,\mathbf{P}(t) + \mathbf{P}(t)\,\mathbf{F}^{\mathsf{T}} + \mathbf{L}\,\mathbf{Q}\,\mathbf{L}^{\mathsf{T}}.$$

- For example, let's consider the spring model:

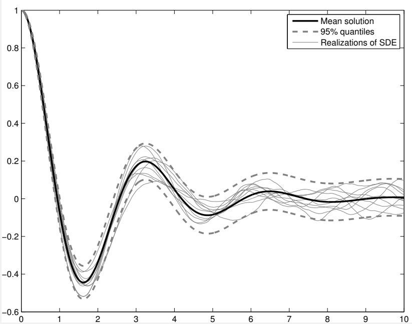
$$\underbrace{\begin{pmatrix} \frac{\mathrm{d}x_1(t)}{\mathrm{d}t} \\ \frac{\mathrm{d}x_2(t)}{\mathrm{d}t} \end{pmatrix}}_{\mathrm{d}\mathbf{x}(t)/\mathrm{d}t} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\nu^2 & -\gamma \end{pmatrix}}_{\mathbf{F}} \underbrace{\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}}_{\mathbf{x}} + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{\mathbf{L}} w(t).$$

# Mean and covariance of LTI SDEs (cont.)

The mean and covariance equations:

$$\begin{pmatrix} \frac{dm_1}{dt} \\ \frac{dm_2}{dt} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\nu^2 & -\gamma \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}$$

$$\begin{pmatrix} \frac{dP_{11}}{dt} & \frac{dP_{12}}{dt} \\ \frac{dP_{21}}{dt} & \frac{dP_{22}}{dt} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\nu^2 & -\gamma \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} + \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -\nu^2 & -\gamma \end{pmatrix}^{\mathsf{T}} + \begin{pmatrix} 0 & 0 \\ 0 & q \end{pmatrix}$$



# Alternative derivation of mean and covariance

- We can also attempt to derive mean and covariance equations directly from

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t), \qquad \mathbf{x}(t_0) \sim \mathrm{N}(\mathbf{m}_0, \mathbf{P}_0).$$

- By taking expectations from both sides gives

$$\mathrm{E}\left[\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}\right] = \frac{\mathrm{d}\,\mathrm{E}[\mathbf{x}(t)]}{\mathrm{d}t} = \mathrm{E}\left[\mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t)\right] = \mathbf{F}\;\mathrm{E}[\mathbf{x}(t)].$$

- This thus gives the correct mean differential equation

$$\frac{\mathrm{d}\mathbf{m}(t)}{\mathrm{d}t} = \mathbf{F}\,\mathbf{m}(t)$$

# Alternative derivation of mean and covariance (cont.)

- For the covariance we use

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[(\mathbf{x}-\mathbf{m})\,(\mathbf{x}-\mathbf{m})^{\mathsf{T}}\right] = \left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}-\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t}\right)(\mathbf{x}-\mathbf{m})^{\mathsf{T}} + (\mathbf{x}-\mathbf{m})\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}-\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t}\right)^{\mathsf{T}}$$

- Substitute $\mathrm{d}\mathbf{x}(t)/\mathrm{d}t = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t)$ and take expectation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathsf{E}\left[(\mathbf{x}-\mathbf{m})\,(\mathbf{x}-\mathbf{m})^{\mathsf{T}}\right] = \mathbf{F}\,\mathsf{E}\left[(\mathbf{x}(t)-\mathbf{m}(t))\,(\mathbf{x}(t)-\mathbf{m}(t))^{\mathsf{T}}\right] + \mathsf{E}\left[(\mathbf{x}(t)-\mathbf{m}(t))\,(\mathbf{x}(t)-\mathbf{m}(t))^{\mathsf{T}}\right]\mathbf{F}^{\mathsf{T}}$$

- This implies the covariance differential equation

$$\frac{\mathrm{d}\mathbf{P}(t)}{\mathrm{d}t} = \mathbf{F}\,\mathbf{P}(t) + \mathbf{P}(t)\,\mathbf{F}^{\mathsf{T}}.$$

- But this solution is wrong!

# Alternative derivation of mean and covariance (cont.)

- Our mistake was to assume

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[(\mathbf{x}-\mathbf{m})\,(\mathbf{x}-\mathbf{m})^{\mathsf{T}}\right] = \left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}-\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t}\right)(\mathbf{x}-\mathbf{m})^{\mathsf{T}} + (\mathbf{x}-\mathbf{m})\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}-\frac{\mathrm{d}\mathbf{m}}{\mathrm{d}t}\right)^{\mathsf{T}}$$

- However, this result from basic calculus is not valid when $\mathbf{x}(t)$ is stochastic.
- The mean equation was ok, because its derivation did not involve the usage of chain rule (or product rule) above.
- But which results are right and which wrong?
- We need to develop a whole new calculus to deal with this...

# Fourier domain solution of SDE

- Consider the scalar SDE (Ornstein–Uhlenbeck process):

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -\lambda\, x(t) + w(t)$$

- Let's take a formal Fourier transform (*Warning: $w(t)$ is not a square-integrable function!*):

$$(i\,\omega)\, X(i\,\omega) = -\lambda\, X(i\,\omega) + W(i\,\omega)$$

- Solving for $X(i\,\omega)$ gives

$$X(i\,\omega) = \frac{W(i\,\omega)}{(i\,\omega) + \lambda}$$

- This can be seen to have the transfer function form

$$X(i\,\omega) = H(i\,\omega)\, W(i\,\omega)$$

  where the transfer function is

$$H(i\,\omega) = \frac{1}{(i\,\omega) + \lambda}$$

# Fourier domain solution of SDE (cont.)

- By direct calculation we get

$$h(t) = \mathscr{F}^{-1}[H(i\,\omega)] = \exp(-\lambda\,t)\,u(t),$$

where $u(t)$ is the Heaviside step function.

- The solution can be expressed as convolution, which thus gives

$$\begin{aligned} x(t) &= h(t) * w(t) \\ &= \int_{-\infty}^{\infty} \exp(-\lambda\,(t-\tau))\,u(t-\tau)\,w(\tau)\,\mathrm{d}\tau \\ &= \int_{0}^{t} \exp(-\lambda\,(t-\tau))\,w(\tau)\,\mathrm{d}\tau \end{aligned}$$

provided that $w(t)$ is assumed to be zero for $t < 0$.

- Analogous derivation works for multidimensional LTI SDEs

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t)$$

# Spectral densities and covariance functions

- A useful quantity is the spectral density which is defined as

$$S_x(\omega) = \mathrm{E}[|X(i\,\omega)|^2] = \mathrm{E}[X(i\,\omega)\,X(-i\,\omega)].$$

- What makes it useful is that the stationary-state covariance function is its inverse Fourier transform:

$$C_x(\tau) = \mathrm{E}[x(t)\,x(t+\tau)] = \mathscr{F}^{-1}[S_x(\omega)]$$

- For the Ornstein–Uhlenbeck process we get

$$S_x(\omega) = \frac{\mathrm{E}[|W(i\,\omega)|^2]}{|(i\,\omega)+\lambda|^2} = \frac{q}{\omega^2+\lambda^2},$$

and

$$C(\tau) = \frac{q}{2\lambda}\,\exp(-\lambda\,|\tau|).$$

# Spectral densities and covariance functions (cont.)

- In multidimensional case we have (joint) spectral density matrix:

$$\mathbf{S}_{\mathbf{x}}(\omega) = \mathrm{E}[\mathbf{X}(i\,\omega)\,\mathbf{X}^{\mathsf{T}}(-i\,\omega)],$$

- The joint covariance matrix is its inverse Fourier transform

$$\mathbf{C}_{\mathbf{x}}(\tau) = \mathscr{F}^{-1}[\mathbf{S}_{\mathbf{x}}(\omega)].$$

- For general LTI SDEs

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t),$$

we get

$$\begin{aligned}
\mathbf{S}_{\mathbf{x}}(\omega) &= (\mathbf{F} - (i\,\omega)\,\mathbf{I})^{-1}\;\mathbf{L}\,\mathbf{Q}\,\mathbf{L}^{\mathsf{T}}\;(\mathbf{F} + (i\,\omega)\,\mathbf{I})^{-\mathsf{T}} \\
\mathbf{C}_{\mathbf{x}}(\tau) &= \mathscr{F}^{-1}[(\mathbf{F} - (i\,\omega)\,\mathbf{I})^{-1}\;\mathbf{L}\,\mathbf{Q}\,\mathbf{L}^{\mathsf{T}}\;(\mathbf{F} + (i\,\omega)\,\mathbf{I})^{-\mathsf{T}}].
\end{aligned}$$

# Problem with general solutions

- We could now attempt to analyze non-linear SDEs of the form

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}, t) + \mathbf{L}(\mathbf{x}, t)\, \mathbf{w}(t)$$

- We cannot solve the deterministic case—no possibility for a “leap of faith”.
- We don’t know how to derive the mean and covariance equations.
- What we can do is to simulate by using Euler–Maruyama:

$$\hat{\mathbf{x}}(t_{k+1}) = \hat{\mathbf{x}}(t_k) + \mathbf{f}(\hat{\mathbf{x}}(t_k), t_k)\, \Delta t + \mathbf{L}(\hat{\mathbf{x}}(t_k), t_k)\, \Delta\beta_k,$$

  where $\Delta\beta_k$ is a Gaussian random variable with distribution $\mathrm{N}(\mathbf{0}, \mathbf{Q}\, \Delta t)$.
- Note that the variance is proportional to $\Delta t$, not the standard deviation.

# Problem with general solutions (cont.)

- Picard–Lindelöf theorem can be useful for analyzing existence and uniqueness of ODE solutions. Let's try that for

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, t) + \mathbf{L}(\mathbf{x}, t)\,\mathbf{w}(t)$$

- The basic assumption in the theorem for the right hand side of the differential equation were:
  - Continuity in both arguments.
  - Lipschitz continuity in the first argument.
- But white noise is discontinuous everywhere!
- We need a new existence theory for SDE solutions as well...

# Summary

- Stochastic differential equation (SDE) is an ordinary differential equation (ODE) with a stochastic driving force.
- SDEs arise in various physics and engineering problems.
- Solutions for linear SDEs can be (heuristically) derived in the similar way as for deterministic ODEs.
- We can also compute the mean and covariance of the solutions of a linear SDE.
- Fourier transform solutions to linear time-invariant (LTI) SDEs lead to the useful concepts of spectral density and covariance function.
- The heuristic treatment only works for some analysis of linear SDEs, and for e.g. non-linear equations we need a new theory.
- One way to approximate solution of SDE is to simulate trajectories from it using the Euler–Maruyama method.

# Matlab demonstration

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -\lambda\, x(t) + w(t), \quad x(0) = x_0,$$