Asymptotic approximations
In the next sections we cover different types of asymptotic behavior in SDE’s.

Examples:
1. Approximating systems with a large number of discrete events with a continuous process: Related to central limit theorem: normal approximation for a large number $N$ of realizations

2. Small noise asymptotics and behavior in the state space for: boundary layers in state space

3. Systems with multiple time scales: quasi-steady approximations and stochastic averaging

   Example: Quasi-steady approximations in deterministic dynamics

Michaelis-Menten Model: (deterministic model)

\[
\begin{align*}
S + E \xrightleftharpoons[k_3]{k_1} C & \xrightarrow[k_2]{k_3} P + E \\
\text{substrate} + \text{enzyme} & \rightarrow \text{complex} \rightarrow \text{further reaction to product} + \text{enzyme}
\end{align*}
\]

\[
\begin{align*}
\dot{S} &= k_1 SE - k_2 C - k_3 C \\
\dot{E} &= -k_1 SE + k_3 C \\
\dot{C} &= -k_1 SE + k_2 C + k_3 C \\
\dot{P} &= k_2 C
\end{align*}
\]

$E + C = \text{conserved quantity} = E_0$ \hspace{1cm} Take $(C(0) = 0)$ so we can eliminate $E$

Here the $k_j$ are the reaction rates.
\[
\begin{align*}
\dot{C} &= k_1 S (E_0 - C) - k_3 C - k_2 C \\
\dot{S} &= -k_1 S (E_0 - C) + k_3 C \\
\dot{P} &= k_2 C \\
\end{align*}
\]

linear stability shows that \( C = 0 \), \( S = 0 \) is a stable fixed point. \( \text{(95)} \)

Note that this is a nonlinear system of ODE's - in general it is difficult to get a closed-form expression for systems of this type.

Instead, suppose \( E_0 \ll 1 \) (small amount of enzyme); then also small amount of \( C \).

The notation \( \ll \) is interpreted as “is much less than”, and is usually related to orders of magnitude - that is the initial amount of enzyme \( E_0 \) is small compared with other quantities that are larger (sometimes stated as \( O(1) \) - order one, that is, a constant that is not large or small).

\[ C = E_0 c \] \( \text{(96)} \)

\[ \Rightarrow E_0 \dot{c} = k_1 S (1 - c) E_0 - k_3 E_0 c - k_2 E_0 c \] \( \text{(97)} \)

\[ \dot{S} = -k_1 S (1 - c) E_0 + k_3 E_0 c \] \( \text{(98)} \)

To see the leading order behavior, \( T = E_0 t \) (note this is a short time)

\[ \Rightarrow E_0 c_T = k_1 S_1 (1 - c) - k_3 c - k_2 c \] \( \text{(99)} \)

\[ S_T = -k_1 S (1 - c) + k_3 c \] \( \text{(100)} \)

Leading order: set \( E_0 c_T = 0 \)

\[ \Rightarrow c = \frac{k_1 S}{k_1 S + k_2 + k_3} \] like a steady-state, if \( S \) was a constant \( \text{(101)} \)

\[ S_T = -k_1 S (1 - c) + k_3 c \] \( \text{(102)} \)

Leading order indicates comparing the relative sizes of different terms, under the assumption that \( E_0 \ll 1 \) while other quantities are \( O(1) \).
The system (102) is the quasi-steady approximation: \( S \) is treated like a constant in the equation for \( c \). The solution is compared to the solution of the full system (94) in the figure. Note that even though \( E_0 \) is not very small, the approximation still does well after an initial transient.

What is the physical interpretation: \( c \) changes quickly to adjust to the value of \( S \), i.e. \( S \) is used up quickly in the reaction with \( E \). Meanwhile \( S \) changes slowly, and looks like a constant relative to \( c \).

Figure 5: Realization M-M process, with \( E_0 \ll 1 \). The red line is \( S \), the blue line is \( C \), both obtained from (94). The +'s are obtained for \( c \) from the quasi-steady approximation (102. Here \( E_0 = .2 \).
6 Stochastic multiple scales analysis

Duffing-van der Pol equation

\[ \frac{dx}{dt} = y dt \]  
\[ \frac{dy}{dt} = \left[-\omega^2 x + \beta y - ax^3 - bx^2 y\right] dt + \delta g(x) dw. \]

(103)  
(104)

\[ g(x) = 1 = \text{additive noise} \]
\[ g(x) = x = \text{multiplicative noise} \]

\[ \delta \neq 0, \text{coherence resonance} \]

Deterministic case: \( \delta = 0 \ |\beta| \ll 1, \beta = \epsilon^2 \beta_2 \epsilon \ll 1 \).
\( \beta < 0 \): Slow decay of oscillations, zero state stable
\( \beta > 0 \): zero state unstable, stable limit cycle

Stochastic case: \( \delta \neq 0 \), coherence resonance

Deterministic multi-scale approximation

Modulation equations for amplitudes of oscillations:
\[ A(T) \cos \omega t + B(T) \sin \omega t \]

\( T \) is a “slow” time: \( T = \epsilon^2 t, \epsilon \ll 1 \).

\( \epsilon \) is typically related to the proximity to a transition point and/or a specific frequency \( \omega \).

van der Pol-Duffing equation:

\[ x_{tt} = -\omega^2 x + \epsilon^2 x_t - ax^3 - bx^2 x_t, \quad \epsilon \ll 1 \]

\[
\begin{align*}
   dx & = y \, dt \\
   dy & = \left[-\omega^2 x + \epsilon y - ax^3 - bx^2 y\right] \, dt
\end{align*}
\]

Amplitudes \( A(T) \) and \( B(T) \) may also scale with \( \epsilon \).

“Multi-scale analysis” or “Averaging”

\( x \sim \epsilon A(T) \cos(\omega t + \phi(T)) + \epsilon^2 x_2 + \epsilon^3 x_3 \Rightarrow \)

\[ x_{3tt} + \omega^2 x_3 = x_{1t} - 2x_{1T} - ax_1^3 - bx_1^2 x_{1t} \]

Eliminate resonances in \( x_1, x_3 \) \( \Rightarrow \)

\[
\frac{dA}{dT} = A \left( \frac{\beta}{2} - \frac{b}{8} (A^2 + B^2) \right) - \frac{3a}{8} (A^2 + B^2) B = \psi(T)
\]

Projection onto \( \cos \omega t \), integration treating \( t \) and \( T \) as independent
Stochastic multi-scale approximation:
\[ \hat{x} \sim \epsilon [A(T) \cos \omega t + B(T) \sin \omega t], \quad T = \epsilon^2 t, \quad \hat{y} \sim -\epsilon \omega \left[ A(T) \sin \omega t - B(T) \cos \omega t \right]. \]
Captures stochastic effects on nearly-regular oscillations

\[ dA = \psi_A dT + \sigma_A d\xi_1(T), \quad dB = \psi_B dT + \sigma_B d\xi_2(T), \]

Find drift, diffusion coefficients:
First, using
\[ d\hat{x} = \frac{\partial \hat{x}}{\partial t} dt + \frac{d\hat{x}}{dA} dA + \frac{d\hat{x}}{dB} dB \]
\[ = \hat{y} dt + \epsilon \cos \omega t (\psi_A dT + \sigma_A d\xi_1) + \epsilon \sin \omega t (\psi_B dT + \sigma_B d\xi_2) \]
\[ d\hat{y} = \frac{\partial \hat{y}}{\partial t} dt + \frac{d\hat{y}}{dA} dA + \frac{d\hat{y}}{dB} dB \]
\[ = [-\omega^2 \hat{x}] dt - \epsilon \omega \sin \omega t (\psi_A dT + \sigma_A d\xi_1) + \epsilon \omega \cos \omega t (\psi_B dT + \sigma_B d\xi_2). \]

Noting \( \frac{d^2\hat{x}}{dA^2} = 0 \) and \( \frac{d^2\hat{y}}{dB^2} = 0 \), Substitution:
\[ d\hat{x} = [-\epsilon (A(T) \sin \omega t + B(T) \cos \omega t)] dt \]
\[ d\hat{y} = \epsilon \left[ -\omega^2 (A(T) \cos \omega t + B(T) \sin \omega t) - \epsilon^2 \omega \beta_2 (A(T) \sin \omega t - B(T) \cos \omega t) \right. \]
\[ -a \epsilon^2 (A(T) \cos \omega t + B(T) \sin \omega t)^3 \]
\[ -b \omega \epsilon^2 (A(T) \cos \omega t + B(T) \sin \omega t)^2 (-A(T) \sin \omega t + B(T) \cos \omega t)] dt + \delta dw. \]

Comparing drift and diffusion coefficients at different orders of \( \epsilon \), we find that the \( O(\epsilon) \) terms cancel.

Then the next terms are \( O(\epsilon^3) \), and for convenience we write them in terms of the slow time \( T \).
\[ (\cos \omega t \psi_A + \sin \omega t \psi_B) dT + \cos \omega t \sigma_A d\xi_1 + \sin \omega t \sigma_B d\xi_2 = 0 \quad (105) \]
\[ \epsilon (-\sin \omega t \psi_A + \cos \omega t \psi_B) dT + \epsilon (-\sin \omega t \sigma_A d\xi_1 + \cos \omega t \sigma_B d\xi_2) = \]
\[ \epsilon [-\beta_2 (A(T) \sin \omega t - B(T) \cos \omega t) - a (A(T) \cos \omega t + B(T) \sin \omega t)^3 \]
\[ -b (A(T) \cos \omega t + B(T) \sin \omega t)^2 (-A(T) \sin \omega t + B(T) \cos \omega t)] dT + \delta dw. \quad (106) \]
Here we have used $\omega = 1$ in the coefficients. Then we equate the drift and diffusion terms. For the diffusion terms we have

\begin{align*}
\cos \omega t \sigma_A d\xi_1 + \sin \omega t \sigma_B d\xi_2 &= 0 \quad (107) \\
\epsilon (-\sin \omega t \sigma_A d\xi_1 + \cos \omega t \sigma_B d\xi_2) &= \delta dw \quad (108) \\
&= \frac{\delta}{\epsilon} \left[ \cos \omega t dw_1(T) + \sin \omega t dw_2(T) \right]
\end{align*}

Here we have used well-known identities:

\[ w(t) = \cos \omega t \, dw_1(t) + \sin \omega t \, dw_2(t) \]

\[ w_j(t) = \epsilon^{-1} w_j(\epsilon^2 t) = \epsilon^{-1} w_j(T) \]

Multi-scale assumption: Fast time $t$ and the slow time scale $T$ are treated as independent for $\epsilon \ll 1$, integrating over one period $2\pi/\omega$

Analogous to deterministic case:
Using orthogonality of $\sin \omega t$ and $\cos \omega t$, and treating $t$ and $T$ as independent, project to get equation on the slow time $T$. For example,

\[ \int_0^{2\pi/\omega} [\epsilon \cos \omega t \cdot (107) - \epsilon \sin \omega t \cdot (108)] \, dt \quad \Rightarrow \sigma_A = -\frac{\delta}{2\epsilon^2}. \quad (109) \]

Similarly, $\sigma_B = \frac{\delta}{2\epsilon^2}$. Then we have the effective noise for the slowly varying stochastic amplitude.

We use the same method to solve for the drift coefficients, substituting the drift terms from (105) and (106) for (107) and (108), respectively, into (109). Again treating functions of $T$ as independent from $t$ yields

\[
\psi_A = \frac{1}{2} \beta_2 A + \frac{3}{8} a B (A^2 + B^2) - \frac{1}{8} b A (A^2 + B^2),
\]

\[
\psi_B = \frac{1}{2} \beta_2 B - \frac{3}{8} a A (A^2 + B^2) - \frac{1}{8} b B (A^2 + B^2),
\]
Stochastic systems with multiple time scales:

\[
\frac{dx}{dt} = f(x, y) \\
\frac{dy}{dt} = \epsilon^{-1} g(x, y)
\]

Can view time scale of \( y \) as “fast”, i.e. for \( T = t/\epsilon \), \( y'(T) = g(x, y) \).

“Slow” time scale \( t \) and “Fast” time scale \( T \)

If fluctuations in \( y \) are fast, then they may appear as fast random fluctuations in the equation for \( x \). Then it may be possible to approximate \( x \) over the slow time scale of \( t \) with

\[
\frac{dX}{dt} = \hat{f}(X, \zeta)
\]  \hspace{1cm} (110)

where \( \zeta \) is a quantity used to approximation the behavior of \( y \). Then \( X \) may approximate the behavior of \( x \) in the weak sense, e.g. in terms of moments or distribution.

How do we determine a reasonable approximation? Under what conditions does this approximation hold?

References:
Mathematical (and applications): Stuart and Pavliotis, Multiscale methods
We show here results for the following system of SDE’s:

\[
\begin{align*}
\frac{dx}{dt} &= f(x, y) dt \\
\frac{dy}{dt} &= \epsilon^{-1} g(x, y) dt + \epsilon^{-1/2} \sigma dW
\end{align*}
\]  

(111)

Specific example:

\[
\begin{align*}
f(x, y) &= -x - ay \\
g(x, y) &= cx - y
\end{align*}
\]  

(112)

Deterministic approximation: Averaged approximation = (A) approximation

Assuming \( y \) is “well-mixed”: that is, on the \( t \) time scale, the density of \( y \) is well-sampled, approaching the stationary density

Then the approximation is based on using the stationary conditional density, \( p(y|x) \). That is, on the fast time scale on which \( y \) fluctuates, \( x \) appears as a constant. Then the proposed (A) approximation is

\[
\frac{dx}{dt} = \bar{f}(\bar{x}) = \int f(x, y)p(y|x) \, dy = E_{y|x}[f(x, y)]
\]  

(113)

For our example, what is \( \bar{f} \)?

If \( x \) is treated as a constant in the equation for \( y \), then \( y \) is an Ornstein-Uhlenbeck (OU) process, with a Gaussian stationary density \( N(cx, \sigma^2/2) \). So

\[
E_{y|x}[f(x, y)] = -x - E[y|x] = -(1 + ac)x \quad \Rightarrow \quad \frac{d\bar{x}}{dt} = -(1 + ac)\bar{x}
\]  

(114)

What are the circumstances for which (A) is a reasonable approximation? Can we do better? Consider our example:

Taking \( x = \bar{x} + \xi \), assuming \( \xi \ll 1 \) is a small correction
\[ \frac{\, d\bar{x}}{\, dt} + \frac{\, d\xi}{\, dt} = f(x, y) - \bar{f}(x) + \bar{f}(x) \quad \Rightarrow \]
\[ \frac{\, d\xi}{\, dt} \sim (f(x, y) - \bar{f}(x)) + \bar{f}(\bar{x}) + \bar{f}'(\bar{x})\xi - \bar{f}(\bar{x}) \]
where we have used a Taylor series about \( \xi = 0 \), and \( \hat{f}(x, y) = (f(x, y) - \bar{f}(x)) \).

Then, to complete the approximation, we need to approximate \( \hat{f}(x, y) \). This term represents the fluctuations of \( f \) about its average, due to fast fluctuations in \( y \). As we look for a weak approximation to \( x \) on the time scale \( t \) (slow time scale), then we look for an approximation to \( \hat{f} \) which has the same properties - in the weak sense.

Specifically, can \( \hat{f} \) be approximated by Brownian motion, or some other random process?

What is the motivation for this approximation?

We expect the properties of \( \hat{f} \) gives the fast fluctuations of \( f(x, y) \) about its average \( \bar{f} \). So if varying on a fast time scale, these can appear as fluctuations of a random process with mean zero. So if we can approximate with something known, then we have a lower dimensional approximation for \( \xi \) and thus for \( x \sim X = \bar{x} + \xi \).

If \( \hat{f} \) has the properties of a Brownian motion \( DdW \), then we consider the integrated behavior of the correlation \( C \):
\[ \int C_{\hat{f}\hat{f}}(\tau) \, d\tau = \int E[\hat{f}(x, y(t + \tau))\hat{f}(x, y(t))] \, d\tau = D^2 \int \delta(\tau) \, d\tau \quad (115) \]
Here \( \tau \) is the lag variable. Usually \( C(\tau) \to 0 \) as \( \tau \to \infty \). That is, correlation between different
points in a realization decreases with increasing time intervals between the points.

Note: for simplicity we take the stationary result for $C_{\hat{f}\hat{f}}$ - that is, as $\epsilon \to 0$, on the fast time scale $t/\epsilon \to \infty$. So $C$ is independent of $t$.
This result can be expressed more generally in the case that the behavior is not identically a Brownian motion, but rather one that in the limit as $\epsilon \to 0$ (time scales are separated) has the behavior of a $\delta$-function.

\[
\int C_{\hat{f}\hat{f}}(\tau)\,dt = \beta \int \frac{1}{\epsilon} h(u/\epsilon)\,du \quad \text{where}
\lim_{\epsilon \to 0} \int \frac{1}{\epsilon} h(u/\epsilon)\,du = \int \delta(u)\,du = 1 \tag{116}
\]

For example, if $\hat{f}$ has the properties of an Ornstein-Uhlenbeck process, on the fast scale, e.g.

\[
dZ = -\frac{\mu}{\epsilon}Z\,dt + \frac{\Sigma}{\sqrt{\epsilon}}\,dW(t) \tag{117}
\]

Then $C_{ZZ}(\tau) = \epsilon\Sigma^2/(2\mu)^2 e^{-\mu|\tau|/\epsilon}$, and

\[
\lim_{\epsilon \to 0} \int \int C_{ZZ}(\tau)d\tau = \epsilon \frac{\Sigma^2}{\mu^2} \tag{118}
\]

Now let’s see how this works for our example:

Recall, $\overline{f}(x) = -(1 + ac)x$

Then $\hat{f} = f(x, y) - \overline{f}(x) = -x - ay + x + acx = -ay + cx$

and $C_{\hat{f}\hat{f}} = a^2E_{y|x}[(y(t) - cx)(y(t + \tau) - cx)] = a^2C_{yy}(\tau)$. Considering the stationary behavior of $y$ (for large time), we have
\[
\int C_{\hat{f}}(\tau) d\tau = \epsilon^2 a^2 \sigma^2 \int_0^\infty \frac{1}{\epsilon} e^{-|\tau|/\epsilon} d\tau = a^2 \sigma^2 \epsilon
\]

Then we conclude that \( \hat{f} \) can be weakly approximated by increments of a Brownian motion \( \epsilon ab dW(t) \), and the slow dynamics for \( x \) in

\[
\begin{align*}
    dx &= -x - ay \\
    dy &= -\frac{cx + y}{\epsilon} + \frac{\sigma}{\sqrt{\epsilon}} dW(t)
\end{align*}
\]

can be weakly approximated as \( x = \bar{x} + \xi \) where

\[
\begin{align*}
    d\bar{x} &= -(1 - ac)\bar{x} \\
    d\xi &= \tilde{f}'(x)\xi + \sqrt{\epsilon}a\sigma dW(t) = -(1 - ac)\xi + \sqrt{\epsilon}a\sigma dW(t)
\end{align*}
\]

This is known as the (L) approximation. It is a linear SDE approximation for the slow dynamics \( x \). We can see that for certain values of the parameters \( a \) and \( b \), namely where \( ab = \epsilon^\gamma \), for \( \gamma > -1 \), the
stochastic fluctuations are negligible as $\epsilon \to 0$. Then $\xi$ decays to its mean 0, and we are left with the (A) approximation.

This form is relatively simple, given that the original system is linear. Furthermore, terms of the form $\overline{f''(x)}\xi^2/2$ vanish. But in a more general nonlinear system, these terms would not vanish and could be a source of error in the approximation.

**General multiple scale averaging approximation (nonlinear):**

The (N+) approximation:

$$dx = \overline{f(x)} dt + D(x) \circ dW \quad \text{(Stratonovich)} \quad (121)$$

$$= \overline{f(x)} dt + G(x) + D dW \quad \text{(Ito)} \quad (122)$$

Note that $\circ$ indicates the Stratonovich interpretation for the noise increment, and $G(x) = D(x)D'(x)/2$ is the correction between the Ito and Stratonovich interpretations.

Derivation of this result:
Reference: Stuart and Pavliotis, Multiscale Methods, Averaging and Homogenization (Springer)

The approach is based on looking at the operator $L$ corresponding to the generator for the full process. Recall, for $u = E(f(x, y))$,

$$u_t = Lu = a \cdot \nabla u + B\nabla^2 u, \quad B = bb^T/2$$

$$dz = a(z) dt + b(z) dW \quad (123)$$

and then looking at the asymptotic behavior of that operator as $\epsilon \to 0$, and identifying the process that corresponds to the asymptotic approximation to $L$. 
Example:

\[ dx = \frac{v(x)y}{\sqrt{\epsilon}} \]
\[ dy = -\frac{\alpha y}{\epsilon} + \frac{\sqrt{2\epsilon}}{\sqrt{\epsilon}} dW(t) \]  \hspace{1cm} (124)

The equation for \( u \) has the form:

\[ u_t = Lu = \left( L_0 + \frac{1}{\sqrt{\epsilon}}L_1 + \frac{1}{\epsilon}L_2 \right) \]  \hspace{1cm} (125)

Using an asymptotic expansion for \( u = u_0 + \sqrt{\epsilon}u_1 + \epsilon u_2 \), then one can consider the sequence of equations obtained at each order of \( \epsilon \).

After a number of calculations, the equation has the form:

\[ L_0 u_2 = M(u_0) \]  \hspace{1cm} (126)

From the Fredholm Alternative Theorem, for there to be a solution we must have

\[ \int M(u_0)\rho(y) \, dy = 0 \quad \text{since} \quad L_0^*\rho(y) = 0 \]  \hspace{1cm} (127)

where \( \rho(y) \) is the invariant density for the fast variable \( y \). Note that this condition is essentially an averaging over \( y \). From that condition we find

\[ \int M(u_0)\rho(y) \, dy = L_x u_0 \]  \hspace{1cm} (128)

where \( L_x \) is the operator corresponding to the reduced equation in (121), which is the N+ approximation.