# Stochastic Toolkit for Uncertainty Quantification in complex nonlinear systems

Michal Branicki and Andrew Majda

September 23, 2011

## Contents

1 Preliminaries 2
   1.1 Basic Probability concepts ............................................. 2
   1.2 Stochastic Differential Equations (SDE's) ............................... 5
      1.2.1 Langevin equation .................................................. 5
      1.2.2 Itô integral and basics of Itô calculus ............................. 7
      1.2.3 Path-wise solutions and statistics of the linear Langevin equation .................. 8
   1.3 Fokker-Planck equation .................................................. 10

2 Gaussian systems 11
   2.1 Linear ‘two-level’ model with one hidden variable and nonnormal dynamics .................. 11
      2.1.1 Path-wise solutions of the two-level model ......................... 12
      2.1.2 Statistics of the two-level system ................................ 13
      2.1.3 Statistics on the attractor ........................................ 14
      2.1.4 Dynamical regimes in the two-level model with stable dynamics ................. 16
   2.2 Different models with the same marginal statistics ...................... 18

3 Non-Gaussian systems 21
   3.1 Linear scalar model with multiplicative noise ........................... 21
   3.2 Exact path-wise solutions for the system (110) .......................... 21
   3.3 Invariant measures for the system (110) ................................ 22
   3.4 Approach to the equilibrium PDF ........................................ 26
   3.5 Moments of the time-dependent PDFs associated with (110) ............... 28
      3.5.1 Moments of the equilibrium PDFs for constant mean damping and forcing ........ 30
   3.6 Canonical scalar model with cubic nonlinearity and deterministic structural instability .. 32
      3.6.1 Motivation .......................................................... 32
      3.6.2 The model and its statistics ...................................... 32
      3.6.3 Regime study ....................................................... 33

4 Building complex models with regime switching 34
   4.1 Markov jump processes ................................................... 34
      4.1.1 Switching times ...................................................... 36
   4.2 Two models with regime switching behaviour ............................. 37
      4.2.1 Linear complex scalar dynamics with transient instabilities ............. 37
      4.2.2 Spatially extended turbulent system with transient instabilities ......... 39
1 Preliminaries

This chapter is a brief and rapid introduction to the basic concepts of probability theory, stochastic processes and Stochastic Differential equations (SDE’s). The exposition is deliberately ‘applied’ and kept concise. More details can be found in virtually any Book on Probability theory or Stochastic Differential Equations.

1.1 Basic Probability concepts

Consider a nonempty set $\Omega$ with elements $\omega$. In order to correctly define the meaning of probability and, subsequently, a probability space we need a few more concepts defined below. In particular one needs to define a distinguished system of subsets of subsets (events) of $\Omega$ forming the so-called $\sigma$-algebra.

**Definition [$\sigma$-algebra]:** A $\sigma$-algebra is a collection of subsets $\mathcal{A}$ of $\Omega$ with the following properties

(i) $\Omega \in \mathcal{A}$,

(ii) If $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$ (Here $A^c \equiv \Omega - A$ is the complement of $A$),

(iii) For a countable collection of sets $A_1, A_2, \cdots \in \mathcal{A}$

$$\cup_{k=1}^{\infty} A_k \in \mathcal{A}. \quad (1)$$

**Remarks:** We will call a set $A \in \mathcal{A}$ an event; elements of $\Omega$ are called sample points. The first condition can be replaced with $\emptyset \in \mathcal{A}$ because $\mathcal{A}$ is closed to complements (i.e. condition (ii)). Note also that (iii) implies that countable intersections of events are also events (due to (i)-(iii) and de Morgan’s formulas).

**Definition [Probability measure]:** Let $\mathcal{A}$ be the $\sigma$-algebra of subsets of $\Omega$. We call $P : \mathcal{A} \to [0,1]$ a probability measure (or simply probability) provided that

(i) $P(\Omega) = 1,$

(ii) If $A_1, A_2, \ldots$ are disjoint sets then

$$P(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k). \quad (2)$$

**Remarks:**

- For any $A \in \mathcal{A}$, we will call $P(A)$ the probability of the event $A$.
- The normalization $P(\Omega) = 1$ implies $P(\emptyset) = 0$.
- Note that $A \in B$ implies $P(A) \leq P(B)$.
- (i) and (ii) imply $P(\cup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} P(A_k)$

**Definition [Probability space]:** The triple $(\Omega, \mathcal{A}, P)$ is called the probability space.
Random variables and the associated distribution and density functions

Definition [Borel algebra of $\mathbb{R}^n$]: The smallest $\sigma$-algebra of subsets of $\mathbb{R}^n$ containing all of its open sets is called a Borel algebra of $\mathbb{R}^n$.

Definition [Random variable]: Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $\mathcal{B}$ be the collection of all Borel subsets of $\mathbb{R}^n$ (i.e. the Borel algebra). A mapping

$$X : \Omega \to \mathbb{R}^n,$$

is called an $n$-dimensional random variable if for each $B \in \mathcal{B}$ we have

$$X^{-1}(B) \in \mathcal{A}.$$  \hspace{1cm} (4)

That is: $X$ is a random variable if it is $\mathcal{A}$-measurable.

Remark: The standard notational convention which we also use here is to write $X$ instead of $X(\omega)$ with explicit dependence of the random variable on the sample point $\omega$. Also, we denote $P(X^{-1}(B))$ as $P(X \in B)$; i.e., the probability that $X$ is in $B$.

Definition [Distribution function]: The distribution function of the random variable $X : \Omega \ni \omega \rightarrow x \in \mathbb{R}^n$ is the function $F_X : \mathbb{R}^n \rightarrow [0, 1]$ defined by

$$F_X(x) := P(X \leq x) \text{ for all } x \in \mathbb{R}^n,$$  \hspace{1cm} (5)

where the notation “$X \leq x$” should be understood component-wise as $X^1 \leq x^1, X^2 \leq x^2, \ldots, X^n \leq x^n$. If $\{X_1, \ldots, X_m\}$ represents a countable collection of $n$-dimensional random variables, then their joint distribution function is $F_{X_1, \ldots, X_m} : \mathbb{R}^{n \times m} \rightarrow [0, 1]$ given by

$$F_{X_1, \ldots, X_m}(x_1; \ldots; x_m) := P(X_1 \leq x_1, \ldots, X_m \leq x_m) \text{ for all } x_i \in \mathbb{R}^n, i = 1, \ldots, m.$$  \hspace{1cm} (6)

Definition [Density function]: Suppose that $X : \Omega \ni \omega \rightarrow x \in \mathbb{R}^n$ is a random variable and $F_X$ its distribution function. If there exists a nonnegative, (Riemann) integrable function $p : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$F_X(x) = \int_{-\infty}^{x^1} \ldots \int_{-\infty}^{x^n} p(y^1, \ldots, y^n)dy^1 \ldots dy^n,$$  \hspace{1cm} (7)

then $p$ is called the (probability) density function for $X$. Given a countable collection of random variables $\{X_1, \ldots, X_m\}$ if there exists a nonnegative, integrable function $p : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$ such that

$$F_{X_1, \ldots, X_m}(x_1; \ldots; x_m) = \int_{-\infty}^{(x_1)^1} \ldots \int_{-\infty}^{(x_1)^n} \int_{-\infty}^{(x_m)^1} \ldots \int_{-\infty}^{(x_m)^n} p(y_1; \ldots; y_m)dy_1 \ldots dy_m,$$  \hspace{1cm} (8)

then $p$ is called the joint (probability) density function for $\{X_1, \ldots, X_m\}$; here $(x_k)^l$ denotes the $l$-th component of $x_k \in \mathbb{R}^n$.  

3
Mean values/expectations

Given an $n$-dimensional ($n < \infty$) random variable $X$ defined over the probability space $(\Omega, \mathcal{A}, P)$, its expectation is defined as a Lebesque integral with respect to the probability measure $P$. Provided that the density function $p$ for $X$ exists, the moments of $p$ can be computed via the notion of the expectation in terms of Riemann integrals over $\mathbb{R}^n$. In fact, all quantities of interest in probability theory can be computed in $\mathbb{R}^n$ in terms of their probability density functions. In particular, the mean and covariance of a random variable $X$ are given by

\[
\begin{align*}
\text{a)} \quad & E[X] = \int_{\Omega} X(\omega) dP(\omega) = \int_{\mathbb{R}^n} x p(x) dx \equiv \langle x \rangle, \\
\text{b)} \quad & \text{Cov}(X) = E[(X - E[X])(X - E[X])^T] = \int_{\mathbb{R}^n} (x - \langle x \rangle)(x - \langle x \rangle)^T p(x) dx, \\
\text{c)} \quad & \text{Var}(X) = \text{diag}(\text{Cov}(X)).
\end{align*}
\]

Clearly, the equalities in (9a,b) assume existence of the probability density, $p$, for $X$.

One of the most widely used distribution is Gaussian (Normal) distribution which is characterized by the following density

- scalar process $X(\omega) = x \in \mathbb{R}$
  \[
p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},
\]
  where $\mu$ and $\sigma^2$ are mean and variance

- in a general case for a vector-valued Gaussian process $X(\omega) = x \in \mathbb{R}^n$ we have the corresponding density
  \[
p(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)},
\]
  where $\mu = \langle x \rangle$ is the mean and $\Sigma = ((x - \mu)(x - \mu)^T)$ is the covariance matrix.

The Normal distribution is usually denoted as $N(\mu, \Sigma)$ with mean $\mu$ and covariance $\Sigma$.

Stochastic processes

Definitions:

(i) A family $\{X_t(\omega) \mid \mathbb{R} \ni t \geq 0\}$ of random variables is called a stochastic process with continuous time

(ii) For each sample point $\omega \in \Omega$, the mapping $t \mapsto X_t(\omega)$ is the corresponding sample path of the stochastic process

Loosely speaking, a stochastic process $X_t(\omega)$ is a ‘random’ function of time, i.e., for each fixed time, $t^*$, $X_{t^*}(\omega)$ is a random variable, and for each fixed sample point, $\omega^*$, $X_t(\omega^*)$ is just a function of time. If we run an experiment and observe the random values of $X_t(\cdot)$ as time evolves we look at a sample path for some fixed $\omega^* \in \Omega$; reruning the same experiment results in observing a different sample path.

The following stochastic process, called Wiener process $W(t)$ or Brownian motions (in physics), is a commonly used stochastic process that satisfies the following properties

Definition [Wiener process (Brownian motion)]: A real-valued stochastic process $W(\cdot)$ is called a Wiener process or Brownian motion if
\begin{itemize}
  \item $W(0) = 0$ with probability one (P-a.s.)
  \item $W(t)$ is continuous P-a.s.
  \item $W(t)$ has independent increments with distribution $W(t) - W(s) \sim N(0, t - s)$ for $0 \leq s < t$
\end{itemize}

Note in particular that
\[ E(W(t)) = 0, \quad E(W^2(t)) = t \quad \text{for each time} \quad t \geq 0. \tag{12} \]

The sample paths of a Wiener process are with probability one nowhere differentiable (they are with probability one nowhere Holder continuous with exponent greater than 1/2). One can also define complex Wiener process as the stochastic process with the independent real and imaginary parts which are real Wiener processes. We also put a factor of $1/\sqrt{2}$ in order to keep the variance of the complex Wiener process the same as in a real case:
\[ W_c(t) = \frac{1}{\sqrt{2}}(W_R(t) + iW_I(t)). \tag{13} \]

We will show a few more example of common stochastic processes through the solutions of SDEs later on.

An important class of stochastic processes which we are concerned with throughout this course consists of Markov processes in which, roughly speaking, the knowledge of only the present state determines the future. Instead of giving here a mathematical definition for (continuous time) stochastic processes, we formulate the Markov assumption in terms of conditional probabilities for stochastic sequences which is sufficient for our purposes.

**Definition** [Markov property (of stochastic sequences)]: Consider a joint probability density, $p(x_1; x_2; \ldots)$, of a stochastic sequence obtained from the process $X$ by sampling it at an ordered sequence of times $t_1 > t_2 > t_3 \ldots$. The process with the density $p$ has the Markov property if the conditional density
\[ p(x_1, t_1; x_2, t_2; \ldots | y_1, \tau_1; y_2, \tau_2; \ldots) = p(x_1, t_1; x_2, t_2; \ldots | y_1, \tau_1), \tag{14} \]
which implies that the conditional probability density associated with the process $X$ is determined entirely by the most recent initial condition $y_1$ at $\tau_1$.

**Remarks:** Wiener process is a Markov process.

### 1.2 Stochastic Differential Equations (SDE’s)

#### 1.2.1 Langevin equation

In order to include the possibility of rapidly varying unresolved dynamics affecting the resolved dynamical system, it seems reasonable to modify the ODE setup as follows
\[ \dot{X}(t) = A(X(t), t) + B(X(t), t)\xi(t), \quad t > 0, \quad X(0) = X_0, \tag{15} \]
where $B : \mathbb{R}^n \to M^{m \times m}$ and $\xi$ is $m$-dimensional ‘white noise’ which is a mathematical idealized concept of a rapidly varying, highly irregular function which requires the existence of a stochastic process with the following properties
\begin{enumerate}
  \item for $t \neq t'$ $\xi(t)$ and $\xi(t')$ are statistically independent, i.e. $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$,
  \item $\xi$ is stationary i.e., the joint distribution $\{\xi(t_1 + t), \ldots, \xi(t_k + t)\}$ does not depend on $t$,
  \item $\langle \xi \rangle = 0$.
\end{enumerate}
Unfortunately, it turns out that any stochastic process satisfying (i) and (ii) cannot have continuous paths which leads to additional complications when defining what it means for the process with continuous sample paths to satisfy a stochastic differential equation of the type (15). Instead of embarking on complicated derivations (see the more ‘pure’ books on SDE’s or Oksendal [9]), one can rewrite the equation (15) in a form that suggests a replacement of $\xi$ with a proper stochastic process. Before proceeding we introduce, in a simplified way, an important concept which is necessary for further derivations:

**Definition** [Nonanticipating function (simplified version)] A function $G(t)$ is called nonanticipating function of $t$ if for all $s$ and $t$ such that $t < s$ $G(t)$ is statistically independent of the increments $W(s) - W(t)$.

Apart from being a reasonable requirement for a solution of a physical process, nonanticipativity of a stochastic process is required for a consistent definition of SDEs in which, loosely speaking, some kind of causality is expected in the sense that unknown future cannot affect the present.

Let now $0 = t_0 < t_1 < \cdots < t_m$ and consider a discrete version of (15)

$$X_{k+1} - X_k = A(X_k, t_k)\Delta t_k + B(X_k, t_k)\xi_k \Delta t_k$$

where

$$X_j = X(t_j), \xi_k = \xi_{t_k}, \Delta t_k = t_{k+1} - t_k$$

We replace $\xi \Delta t_k$ with $\Delta W_k$ where $\{W_k\}$ is some suitable stochastic process. The requirements (i)-(iii) imply that such a process should have stationary independent increments with mean zero. Recall that the only such process with continuous paths is the Wiener process $W(t)$. Thus we put

$$X_{k+1} - X_k = A(X_k, t_k)\Delta t_k + B(X_k, t_k)\Delta W_k$$

It can be shown that the limit of the rhs exists in some appropriate sense for which we adopt the usual integration notation,

$$X(t) = \int_0^t A(X(s), s)ds + \int_0^t B(X(s), s)dW(s)$$

with the meaning of the second integral yet to be defined, and assume that (19) really means that $X(t, \omega)$ is a stochastic process satisfying an SDE which we write as

$$dX(t) = A(X(t), t)dt + B(X(t), t)dW(t), \quad t > 0, X(0) = X_0.$$ (20)

As will soon become clear this interpretation consistent with the physical assumptions will require a rather different definition of the stochastic integral with some unusual properties.

**A note on the white noise process**

A frequently used notational convention is to write $\xi(t) = \dot{W}(t)$ and use the ODE-like notation for an SDE; however, one has to be careful with the interpretation of this notation since the Wiener process is a.s. nowhere differentiable, as pointed out earlier. Thus, $W$ denotes in this context a generalized stochastic process called white noise process which is constructed as a probability measure on the space of tempered distributions on $[0, \infty)$. White noise can be also represented as a generalized mean-squares derivative of the Wiener process. The name white noise comes from the requirement that its autocorrelation function satisfies

$$G(\tau) \equiv \langle \xi(t)\xi(t+\tau) \rangle = \delta(\tau)$$

Consequently, due to the Wiener-Khinchin theorem, the spectral density of a white process is flat since

$$S(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda \tau} G(\tau) d\tau = \frac{1}{2\pi},$$

similar to the spectrum of idealized ‘white’ light.
1.2.2 Ito integral and basics of Ito calculus

Definition [Ito stochastic integral] Given a nonanticipating function \(G(x,t) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^m\) the Ito stochastic integral is defined by

\[
\int_{t_0}^{t} G(x(t'), t')dW(t') = \lim_{n \rightarrow \infty} \sum_{j=1}^{n} G(x(t_{j-1}), t_{j-1}) (W(t_j) - W(t_{j-1})) \tag{23}
\]

where \(x\) is a stochastic process and the limit is taken in the mean square sense (i.e. \(m.s. \lim_{n \rightarrow \infty} \langle (X_n - X)\rangle = 0\)).

The most important properties of Ito calculus are (assuming that \(G, H\) are nonanticipating; the Wiener process, \(W\), is nonanticipating by definition):

- \(\int_{t_0}^{t} G(t')dW(t') = \int_{t_0}^{t} G(t')dt', \) (common abbreviation of this property is \(dW^2(t) = dt\)),
- \(\left\langle \int_{t_0}^{t} G(t')dW(t') \right\rangle = 0,\)
- \(\left\langle \int_{t_0}^{t} G(t')dW(t') \int_{t_0}^{t} H(t')dW(t') \right\rangle = \int_{t_0}^{t} \langle G(t')H(t') \rangle dt',\)
- \(df(W(t), t) = \left( \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} \right) dt + \frac{\partial f}{\partial x} dW(t).\)

Based on the above properties and the properties of the Wiener process it can be deduced that

\[
\int_{t_0}^{t} W(t')dW(t') = \frac{1}{2} \left[ W^2(t) - W^2(t_0) - (t - t_0) \right] \quad \text{and} \quad \left\langle \int_{t_0}^{t} W(t')dW(t') \right\rangle = 0. \tag{24}
\]

This important property of the Ito stochastic integral also points to fundamental differences with regard to the Lebesgue integration (see e.g. Gardiner 2010 for more details.).

Ito formula

Consider now a scalar stochastic process \(X\) satisfying (20). It can be shown, using the above properties of the Ito integral and the fact that \(X\) satisfies (20), that for a smooth deterministic function \(f\) we have

\[
df(X(t)) = \left[ a(X, t)f'(X) + \frac{1}{2} b^2(X, t)f''(X) \right] dt + b(X, t)f'(X)dW(t), \tag{25}
\]

where we skipped the explicit dependence on time in \(X\) for clarity. Analogously, for an \(n\)-dimensional process \(\mathbf{X}\) we have

\[
df(\mathbf{X}(t)) = \left[ \sum_i A_i(\mathbf{X}, t) \partial_i f(\mathbf{X}) + \frac{1}{2} \sum_i [B(\mathbf{X}, t) B^T(\mathbf{X}, t)]_{ij} \partial_i \partial_j f(\mathbf{X}) \right] dt + \sum_{ij} B_{ij}(\mathbf{X}, t) \partial_i f(\mathbf{X}) dW_j(t) \tag{26}
\]

where \(A \in \mathbb{R}^n, B \in \mathbb{R}^{n \times n}\) and \(W\) is an \(m\)-dimensional Wiener process (see, e.g., Gardiner 2010 for the simplest explanation).
1.2.3 Path-wise solutions and statistics of the linear Langevin equation

In order to illustrate some standard solution methods for the sample paths and the second order statistics, we consider linear Langevin equation associated with the so-called Ornstein-Uhlenbeck process

\[ du(t) = [Au(t) + F(t)]dt + B(t)dW(t), \]  

(27)

where \( u(t) \) is an \( N \) dimensional vector, \( A \) is an \( N \times N \) matrix, \( B \) is an \( N \times K \) matrix and \( W \) is \( K \) dimensional Wiener process. The solution \( u(t) \) that satisfies this SDE is the multivariate time dependent forced Ornstein-Uhlenbeck (OU) process. Consider first the homogeneous part of (27) resulting in the following ODE

\[ \frac{du(t)}{dt} = Au(t), \]  

(28)

with solutions

\[ u(t) = e^{A(t-t_0)}u(t_0) \]  

(29)

This solution suggests the following change of variables

\[ v(t) = e^{-A(t-t_0)}u(t). \]  

(30)

so that the equation (27) becomes

\[ dv(t) = e^{-A(t-t_0)}F(t) + e^{-A(t-t_0)}B(t)dW(t). \]  

(31)

According to (19) this equation is solved by

\[ v(t) = v(t_0) + \int_{t_0}^{t} e^{-A(s-t_0)}F(s)ds + \int_{t_0}^{t} e^{-A(s-t_0)}B(s)dW(s). \]  

(32)

where the first integral is the usual Stieltjes integral and the integration in the third term is carried out in the Ito sense (23). Returning to the original variables we obtain

\[ u(t) = e^{A(t-t_0)}u(t_0) + \int_{t_0}^{t} e^{A(t-s)}F(s)ds + \int_{t_0}^{t} e^{A(t-s)}B(s)dW(s). \]  

(33)

Note, that \( u(t) \) stays Gaussian given Gaussian initial data. Therefore, the pdf of \( u(t) \) is fully determined by its first two moments.

**Statistics**

First of all note that, the process \( u(t) \) given by (33) is Gaussian since, due to the properties of the Ito stochastic integral, the last term in (33) is Gaussian. Using the tools introduced in the previous section we can easily determine the second-order statistics of the solutions (33) which due to the Gaussianity of \( u(t) \) completely describes its probability density.

The mean solution is given by

\[ \bar{u}(t) \equiv \langle u(t) \rangle = e^{A(t-t_0)}u(t_0) + \int_{t_0}^{t} e^{A(t-s)}F(s)ds. \]  

(34)
The covariance is, by definition,
\[
R(t) \equiv \langle (u(t) - \langle u(t) \rangle)(u(t) - \langle u(t) \rangle)^T \rangle = e^{A(t-t_0)} R(t_0) e^{A^T(t-t_0)} + \int_{t_0}^{t} e^{A(t-s)} B(s) B^T(s) e^{A^T(t-s)} ds. \tag{35}
\]

Finally, the autocorrelation of \( u(t) \) is given by
\[
Corr(u(t), u(s)) = \langle u(t)u(s) \rangle - \langle u(t) \rangle \langle u(s) \rangle
= e^{A(t-t_0)} Corr(u(t_0), u(t_0)) e^{A^T(t-t_0)} + \int_{t_0}^{\min(t,s)} e^{A(t-t')} B(t') B^T(t') e^{A^T(t-t')} dt'.
\]

Now, we consider the unforced one-dimensional version of (27) with constant coefficients, i.e., \( F(t) = 0, A(t) = -\gamma, B(t) = \sigma \). Then, the solution becomes
\[
u(t) = u(t_0) e^{-\gamma(t-t_0)} + \sigma \int_{t_0}^{t} e^{-\gamma(t-t')} dW(t'). \tag{36}\]

We easily compute mean and variance of \( u(t) \).
\[
\langle u(t) \rangle = \langle u(t_0) \rangle e^{-\gamma(t-t_0)},
Var(u(t)) = Var(u(t_0)) e^{-2\gamma(t-t_0)} + \frac{\sigma^2}{2\gamma} \left( 1 - e^{-2\gamma(t-t_0)} \right).
\]

When the system reaches equilibrium, we find that
\[
\langle u(t) \rangle_{eq} = 0,
Var_{eq}(u(t)) = \frac{\sigma^2}{2\gamma}.
\]

Since \( u(t) \) is Gaussian, then in the equilibrium state we find the following equilibrium pdf of the values of \( u(t) \)
\[
p_{eq}(u) = \frac{1}{\sqrt{2\pi \sigma^2 / 2\gamma}} e^{-\frac{u^2}{2 \sigma^2 / 2\gamma}}. \tag{37}\]

Next, we calculate the time auto-correlation function
\[
Corr(u(t), u(s)) = \langle u(t)u(s) \rangle - \langle u(t) \rangle \langle u(s) \rangle
\]
\[
= Var(u(t_0)) e^{-\gamma(t+s-2t_0)} + \sigma^2 \left( \int_{t_0}^{t} e^{-\gamma(t-t')} dW(t') \int_{t_0}^{t} e^{-\gamma(t-t')} dW(t') \right)
\]
\[
= Var(u(t_0)) e^{-\gamma(t+s-2t_0)} + \sigma^2 \int_{t_0}^{\min(t,s)} e^{-\gamma(t+s-2t')} dt'
\]
\[
= \left( Var(u(t_0)) - \frac{\sigma^2}{2\gamma} \right) e^{-\gamma(t+s-2t_0)} + \frac{\sigma^2}{2\gamma} e^{-\gamma|t-s|}. \tag{38}\]

Therefore, in the equilibrium state, the time auto-correlation function for the process \( u(t) \) satisfying the one-dimensional version of (27) with constant coefficients becomes
\[
Corr_{eq}(\tau) = \frac{\sigma^2}{2\gamma} e^{-\gamma \tau}. \tag{39}\]
1.3 Fokker-Planck equation

Fokker-Planck equation (FPE) describes the evolution of the probability density \( p(x,t) \) associated with a stochastic process \( x(t) \) satisfying an SDE (20). Consider first a scalar process \( X(t) \) satisfying
\[
dX = a(X,t)dt + b(X,t)dW(t).
\]
Consider now any smooth function \( f(x(t)) \). Then the evolution of its mean can be given by the Ito’s formula
\[
dt \langle f(X(t)) \rangle = a(X(t),t) \frac{\partial f}{\partial x} + \frac{1}{2} b(X(t),t)^2 \frac{\partial^2 f}{\partial x^2} \int p(X,t)dx.
\]
On the other hand we have, by definition,
\[
dt \langle f(x(t)) \rangle = \int f(x) \frac{\partial p(x,t)}{\partial t} dx.
\]
Integrating (41) by parts and using the fact that \( f(x(t)) \) was chosen arbitrary we find that the density \( p(x,t) \) satisfies the following PDE
\[
\frac{\partial p}{\partial t} = - \frac{\partial}{\partial x} \left( a(x,t)p(x,t) \right) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left( b(x,t)^2 p(x,t) \right).
\]
with the initial condition \( p \big|_{t=t_0} = p_0 \).

Similarly, for an \( n \)-dimensional process \( x(t) \) \( \in \mathbb{R} \) satisfying
\[
dx = A(x,t)dt + B(x,t)dW(t),
\]
we find that the corresponding FPE for the evolution of the probability density associated with \( x(t) \) is given by
\[
\frac{\partial p}{\partial t} = \mathcal{L}_{FP} p = - \sum_i \frac{\partial}{\partial x_i} \left( A_i(x,t) p \right) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} \left( \left( B(x,t)B^T(x,t) \right)_{ij} p \right),
\]
with the initial condition \( p \big|_{t=t_0} = p_0 \).

The linear operator \( \mathcal{L}_{FP} \) defined in (45) is called the Fokker-Planck operator. In the special case of deterministic dynamics, i.e., when \( \sigma = 0 \), the Fokker-Planck operator reduces to the Liouville operator. The linearity of FPE is deceptive and in many practical situations this equation cannot be solved or reasonably well approximated numerically due to, typically, large number of dimensions involved and problems with imposing the right boundary conditions. On the other hand, stationary solutions in the scalar case are usually easily obtained for a range of linear and nonlinear problems. It is instructive to consider here the stationary solutions of the scalar FPE the linear case which has a Gaussian invariant measure (i.e. equilibrium probability density). Consider the scalar system (40) with
\[
a(x,t) = -\gamma x, \quad b(x,t) = \sigma, \quad \gamma, \sigma = \text{const.} > 0,
\]
so that the associated FPE is
\[
\frac{\partial p}{\partial t} = \gamma \frac{\partial}{\partial x} (xp) + \frac{\sigma^2}{2} \frac{\partial^2 p}{\partial x^2}.
\]
The stationary (equilibrium) solutions of (46) satisfy
\[
\frac{\partial p}{\partial t} = 0.
\]
which implies that
\[ xp + \frac{\sigma^2}{2\gamma} \frac{\partial p}{\partial x} = c_1, \tag{48} \]
for some constant \( c_1 \); however, the boundary conditions \( p \to 0 \) and \( \frac{\partial p}{\partial x} \to 0 \) imply that \( c_1 = 0 \). Thus, the stationary solutions \( p_{eq} \) of (46) are Gaussians given by
\[ p(x) = N_0 \exp \left( -\frac{1}{2} \frac{x^2}{\sigma^2/2\gamma} \right), \tag{49} \]
where \( N_0 \) is a normalizing constant. Note that the above solution coincides with (37) computed earlier directly from the SDEs by finding its first and second moments. It is worth stressing, however, that the methods for finding invariant measures from FPE is much more versatile (at least in the scalar case) and allows for computing non-Gaussian measures.

2 Gaussian systems

We start our discussion by introducing a simple two-dimensional forced linear stochastic model which plays an important role for two reasons:

- it serves as a good pedagogical example allowing for illustration of the previously introduced concepts beyond the scalar setting
- it provides the simplest example of a system with ‘resolved’ dynamics which is affected by the ‘hidden’ unresolved variable. Throughout this course we will be concerned with quantifying consequences of the interaction between the resolved, coarse-grained dynamics and the unresolved scales in complex nonlinear systems. The linear model introduced below allows for illustrating many important concepts in an unambiguous fashion.

2.1 Linear ‘two-level’ model with one hidden variable and nonnormal dynamics

Consider the following \( 2 \times 2 \) linear stochastic system
\[
\begin{align*}
\, dx &= (-ax + cy + F_x(t))dt, \\
\, dy &= (\epsilon^{-1}x - by + F_y(t))dt + \sigma dW(t),
\end{align*}
\tag{50}
\] with parameters \( a, b, \epsilon \) deterministic forcing \( \mathbf{F} = (F_x, F_y) \) and a constant noise amplitude \( \sigma > 0 \); \( W \) is here a scalar Wiener process. It is convenient to rewrite (50) in a more compact form as
\[
\begin{pmatrix}
\, dx \\
\, dy
\end{pmatrix} = \begin{pmatrix}
\hat{L} \begin{pmatrix}
\, dx \\
\, dy
\end{pmatrix} + \begin{pmatrix}
F_x(t) \\
F_y(t)
\end{pmatrix} \\
\, 0
\end{pmatrix} dt + \begin{pmatrix}
0 \\
\sigma
\end{pmatrix} dW, \\
\tag{52}
\end{align*}
\]
where
\[
\hat{L} = \begin{pmatrix}
-a & \epsilon \\
\epsilon^{-1} & -b
\end{pmatrix}, \quad a, b, \epsilon \in \mathbb{R},
\tag{53}
\]
\[
\det(\hat{L}) = \lambda_1 \lambda_2 = ab - 1, \quad \text{tr}(\hat{L}) = \lambda_1 + \lambda_2 = -(a + b),
\tag{54}
\]
and the eigenvalues $\lambda_{1,2}$ are given by
\[
\lambda_1 = -\frac{1}{2} \left( a + b - \sqrt{(a-b)^2 + 4} \right), \quad \lambda_2 = -\frac{1}{2} \left( a + b + \sqrt{(a-b)^2 + 4} \right).
\] (55)

Clearly, in the deterministic case unforced case, i.e. when $\sigma = 0$ and $F \equiv 0$, the system (52) has an attractor provided that
\[
\lambda_1 \lambda_2 > 0, \quad \lambda_1 + \lambda_2 > 0 \Rightarrow ab - 1 > 0, \quad a + b > 0,
\] (56)

which assure that $\hat{L}$ has eigenvalues with negative real parts.

In the remainder of this section we derive analytical path-wise solutions of (52) and its exact second-order statistics in the general case when $\sigma \neq 0$ and show that for any reasonably well-behaved forcing and $0 < \epsilon < \infty$ the system has an attractor provided that (56) hold. We then use these simple analytical results to illustrate the following important points

- Nonnormal behaviour of the system (52), which occurs for $\epsilon \ll 0$ and $\epsilon \gg 1$ when the eigenvectors of $\hat{L}$ are nearly parallel, and its potential consequences in systems with hidden (unresolved) dynamics

- Nonuniqueness of marginal dynamics and the existence of families of systems (50) with the same marginal statistics.

Recall that, as already shown in Section ??, the dynamics of (52) is Gaussian for Gaussian initial conditions $x_0$ and $y_0$; moreover, since (52) is linear the covariance of $(x,y)^T$ is independent of the external forcing.

### 2.1.1 Path-wise solutions of the two-level model

Following the same procedure as in Section 1.2.3, the solutions of the linear system (52) can be easily found in the form
\[
\begin{pmatrix} x \\ y \end{pmatrix} = e^{\hat{L}(t-t_0)} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \int_{t_0}^{t} e^{\hat{L}(t-s)} \begin{pmatrix} F_x(s) \\ F_y(t) \end{pmatrix} ds + \sigma \int_{t_0}^{t} e^{\hat{L}(t-s)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW(s),
\] (57)

where the operator $e^{\hat{L}t}$ is explicitly given by
\[
e^{\hat{L}t} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} (\lambda_1 + a)e^{\lambda_2 t} - (\lambda_2 + a)e^{\lambda_1 t} & \epsilon \left( e^{\lambda_1 t} - e^{\lambda_2 t} \right) \\ \epsilon^{-1} \left( e^{\lambda_1 t} - e^{\lambda_2 t} \right) & (\lambda_1 + a)e^{\lambda_1 t} - (\lambda_2 + a)e^{\lambda_2 t} \end{pmatrix},
\] (58)

with the eigenvalues $\lambda_{1,2}$ of $\hat{L}$ given by (55). Clearly, the conditions (56) imply that for any bounded $F$ and $0 < \epsilon < \infty$ the path-wise solutions (57) with any initial $(x_0,y_0)$ collapse onto the particular solution given by the last two terms in (57). The approach to the attractor depends on the system parameters and in certain parameter regimes the transient dynamics may lead to interesting effects on uncertainty quantification when only $x$ is resolved. As will be shown in Section 2.1.4, these effects depend on both the orientation of the eigenvectors of $\hat{L}$ and the angle between them.

\[
v_1 = \begin{pmatrix} \epsilon / (\lambda_1 + a) \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} \epsilon / (\lambda_2 + a) \\ 1 \end{pmatrix},
\] (59)

and the angle, $\alpha$, between these eigenvectors is
\[
\alpha(\epsilon) = \cos^{-1} \left( \frac{|1 - \epsilon^2|}{\sqrt{\epsilon^4 + \epsilon^2((a-b)^2 + 2) + 1}} \right).
\] (60)
Clearly, the eigenvectors are parallel for $\epsilon \to 0$, and for $\epsilon \to \infty$, and orthogonal for $\epsilon = 1$. We discuss the properties of the system (52) as a function of the parameter $\epsilon$ in Section 2.1.4.

2.1.2 Statistics of the two-level system

Below we derive the exact form of the stochastic attractor. Similarly to the deterministic case, the mean dynamics on the attractor is independent on the initial conditions $\langle x_0, y_0 \rangle$. Since (52) is linear and Gaussian, the stochastic addition to the dynamics on the attractor is fully characterized by the covariance $\Sigma$ and autocorrelation function $C_{\text{att}}$, both of which are independent on the respective initial conditions.

**Mean dynamics of the system (52)**

Similarly to the linear example discussed in §1.2.3, the mean of (57) is given by

$$\langle x(t, t_0) \rangle = e^{L(t-t_0)} \langle x_0 \rangle + \int_{t_0}^{t} e^{L(t-s)} \left( \frac{F_x(s)}{F_y(s)} \right) ds,$$

where we used the properties of the Ito stochastic integral (18). The above formula can be written as

$$\langle x \rangle (t) = \frac{(\lambda_1 + a) e^{\lambda_1(t-t_0)} - (\lambda_2 + a) e^{\lambda_2(t-t_0)}}{\lambda_1 - \lambda_2} \langle x_0 \rangle + \frac{\epsilon (e^{\lambda_1 t} - e^{\lambda_2 t})}{\lambda_1 - \lambda_2} \langle y_0 \rangle$$

$$+ \frac{1}{\lambda_1 - \lambda_2} \int_{t_0}^{t} \left( (\lambda_1 + a) e^{\lambda_1(t-s)} - (\lambda_2 + a) e^{\lambda_2(t-s)} \right) F_x(s) ds,$$

$$\langle y(t) \rangle = \frac{e^{\lambda_1(t-t_0)} - e^{\lambda_2(t-t_0)}}{\epsilon (\lambda_1 - \lambda_2)} \langle x_0 \rangle + \frac{(\lambda_1 + a) e^{\lambda_1(t-t_0)} - (\lambda_2 + a) e^{\lambda_2(t-t_0)}}{\lambda_1 - \lambda_2} \langle y_0 \rangle$$

$$+ \frac{1}{\epsilon (\lambda_1 - \lambda_2)} \int_{t_0}^{t} \left( e^{\lambda_1(t-s)} - e^{\lambda_2(t-s)} \right) F_x(s) ds$$

$$+ \frac{1}{\lambda_1 - \lambda_2} \int_{t_0}^{t} \left( (\lambda_1 + a) e^{\lambda_1(t-s)} - (\lambda_2 + a) e^{\lambda_2(t-s)} \right) F_y(s) ds.$$

**Covariance and autocovariance of $x(t)$ satisfying (52)**

The covariance matrix, $R \equiv \langle (x - \langle x \rangle)(x - \langle x \rangle)^T \rangle$, of the process $x(t, t_0)$ satisfying (52) is given by

$$R(t, t_0) = e^{L(t-t_0)} R_0 e^{L^T(t-t_0)} + \int_{t_0}^{t} e^{L(t-s)} \left( \begin{array}{cc} 0 & 0 \\ 0 & \sigma^2 \end{array} \right) e^{L^T(t-s)}.$$

(64)

The components of $R$ for the system (52) can be easily obtained but the resulting expressions are long and not necessary for our purposes. Thus, we leave these computations as an exercise and only discuss the values of covariance and autocovariance on the statistical attractor below.
2.1.3 Statistics on the attractor

Assuming that the conditions (56) hold so that the eigenvalues \( \lambda_{1,2} \) have negative real parts, one can consider the dynamics of (52) in the limit when the initial conditions are imposed at a very distant past, which is formally obtained by taking the limit \( t_0 \to -\infty \). Consequently, the mean on the attractor is given by

\[
\langle x(t) \rangle_{\text{att}} = \lim_{t_0 \to -\infty} \left[ \frac{1}{\lambda_1 - \lambda_2} \int_{t_0}^{t} \left( (\lambda_1 + a)e^{\lambda_1(t-s)} - (\lambda_2 + a)e^{\lambda_2(t-s)} \right) F_x(s) ds \right. \\
+ \left. \frac{\epsilon}{\lambda_1 - \lambda_2} \int_{t_0}^{t} \left( e^{\lambda_1(t-s)} - e^{\lambda_2(t-s)} \right) F_y(s) ds \right],
\]

(65)

\[
\langle y(t) \rangle_{\text{att}} = \lim_{t_0 \to -\infty} \left[ \frac{1}{\epsilon(\lambda_1 - \lambda_2)} \int_{t_0}^{t} \left( e^{\lambda_1(t-s)} - e^{\lambda_2(t-s)} \right) F_x(s) ds \right. \\
+ \left. \frac{1}{\lambda_1 - \lambda_2} \int_{t_0}^{t} \left( (\lambda_1 + a)e^{\lambda_1(t-s)} - (\lambda_2 + a)e^{\lambda_2(t-s)} \right) F_y(s) ds \right].
\]

(66)

The covariance matrix on the attractor is

\[
\Sigma \equiv \lim_{t_0 \to -\infty} R(t, t_0) = -\frac{\sigma^2}{2(\lambda_1 + \lambda_2)\lambda_1\lambda_2} \begin{pmatrix} \epsilon & \epsilon a \\ \epsilon a & \lambda_1\lambda_2 + a^2 \end{pmatrix}.
\]

(67)

The eigenvectors of this symmetric, positive, semi-definite matrix are

\[
w_1 = \begin{pmatrix} 1 \\ \frac{\lambda_1\lambda_2 + a^2 - \epsilon^2 - \sqrt{(\lambda_1\lambda_2 + a^2 - \epsilon^2)^2 + 4\epsilon^2a^2}}{2\epsilon a} \end{pmatrix},
\]

(68)

\[
w_2 = \begin{pmatrix} \frac{2\epsilon a}{\lambda_1\lambda_2 + a^2 - \epsilon^2 + \sqrt{(\lambda_1\lambda_2 + a^2 - \epsilon^2)^2 + 4\epsilon^2a^2}} \\ 1 \end{pmatrix}, \quad w_1 \cdot w_2 = 0,
\]

(69)

and the corresponding eigenvalues are

\[
\xi_1 = -\frac{\sigma^2}{4(\lambda_1 + \lambda_2)\lambda_1\lambda_2} \left( \lambda_1\lambda_2 + a^2 + \epsilon^2 - \sqrt{(\lambda_1\lambda_2 + a^2 - \epsilon^2)^2 + 4\epsilon^2a^2} \right),
\]

(70)

\[
\xi_2 = -\frac{\sigma^2}{4(\lambda_1 + \lambda_2)\lambda_1\lambda_2} \left( \lambda_1\lambda_2 + a^2 + \epsilon^2 + \sqrt{(\lambda_1\lambda_2 + a^2 - \epsilon^2)^2 + 4\epsilon^2a^2} \right).
\]

(71)

Finally, the autocovariance on the attractor is given by

\[
C_{\text{att}}(\tau, t) = \lim_{t_0 \to -\infty} \left( \begin{pmatrix} x(t, t_0) - \langle x(t, t_0) \rangle \\ y(t, t_0) - \langle y(t, t_0) \rangle \end{pmatrix} \begin{pmatrix} x(t + \tau, t_0) - \langle x(t + \tau, t_0) \rangle \\ y(t + \tau, t_0) - \langle y(t + \tau, t_0) \rangle \end{pmatrix}^T \right) = \Sigma e^{\xi_1 \tau}.
\]

(72)
We now consider two special cases in more detail in order to understand the main dynamical features of these linear systems. In the next section we will discuss the most important role of the model parameters on the system’s dynamics.

Statistics on attractor for constant forcing

For constant forcing, \( F_x(t) = F_x^0, \ F_y(t) = F_y^0, \) the system (52) has a stable invariant measure with equilibrium mean

\[
\langle x \rangle_{eq} = -\frac{(\lambda_1 + \lambda_2 + a)\lambda_1}{\lambda_1\lambda_2}F_x^0 + \frac{\epsilon}{\lambda_1\lambda_2}F_y^0, \tag{73}
\]

\[
\langle y \rangle_{eq} = \frac{1}{\epsilon\lambda_1\lambda_2}F_x^0 + \frac{a}{\lambda_1\lambda_2}F_y^0, \tag{74}
\]

and equilibrium covariance given by (67). This can be easily established by setting \( F_x^1 = F_y^1 = 0 \) in (79) and (82). The equilibrium mean is controlled by \( (\epsilon, F_x^0, F_y^0) \). The covariance \( \Sigma \) and the autocovariance \( C_{att} \) on the attractor are given, respectively, by (67) and (72). Note that both the covariance and autocovariance do not depend on the details of the external forcing for linear systems.

Statistics on attractor for time-periodic forcing

Consider now the system (52) with a simple time-periodic forcing in the form

\[
\begin{align*}
F_x(t) &= F_x^0 + F_x^1 \sin(\omega_x t), \\
F_y(t) &= F_y^0 + F_y^1 \sin(\omega_y t).
\end{align*} \tag{75}
\]

Then, the mean on the attractor (65)-(66) becomes

\[
\langle x(t) \rangle_{att} = \frac{(\lambda_1 + \lambda_2 + a)\lambda_1}{\lambda_1\lambda_2}F_x^0 + \frac{\epsilon}{\lambda_1\lambda_2}F_y^0
\]

\[
+ F_x^1 \left( \omega_x^2 - \lambda_1\lambda_2(\lambda_1 + \lambda_2 + a) \right) \sin \omega_x t - \left( \omega_x^2 + \lambda_1^2 + \lambda_2^2 + \lambda_1\lambda_2 + (\lambda_1 + \lambda_2)a \right) \omega_x \cos \omega_x t \tag{77}
\]

\[
- \frac{\epsilon F_y^1 \left( \omega_y^2 - \lambda_1\lambda_2 \right) \sin \omega_y t - \omega_y(\lambda_1 + \lambda_2) \omega_y \cos \omega_y t}{(\lambda_1^2 + \omega_y^2)(\lambda_2^2 + \omega_y^2)}, \tag{78}
\]

and

\[
\langle y(t) \rangle_{att} = \frac{1}{\epsilon\lambda_1\lambda_2}F_x^0 + \frac{a}{\lambda_1\lambda_2}F_y^0
\]

\[
- F_x^1 \left( \omega_x^2 - \lambda_1\lambda_2 \right) \sin \omega_x t + \left( \lambda_1 + \lambda_2 \right) \omega_x \cos \omega_x t \tag{80}
\]

\[
\frac{\epsilon(\lambda_1^2 + \omega_x^2)(\lambda_2^2 + \omega_x^2)}{\lambda_1\lambda_2}
\]

\[
+ F_y^1 \left( \omega_y^2 - \lambda_2 a \right) \lambda_1 + \omega_y^2(\lambda_2 + a) \right) \sin \omega_y t - \left( (\lambda_2 + a)\lambda_1 - \omega_y^2 + \lambda_2 a \right) \omega_y \cos \omega_y t. \tag{81}
\]

Note again, that the covariance and autocovariance on the attractor are given by (67) and (72) as before and they do no depend on the external forcing. Clearly, the mean on the attractor is controlled in this case by \( (\epsilon, F_x^0, F_x^1, F_y^0, F_y^1, \omega_x, \omega_y) \).
2.1.4 Dynamical regimes in the two-level model with stable dynamics

As already discussed earlier, for $\Re[e^{\lambda_1,2}] < 0$, the system (52) has a global statistical attractor. The approach to the attractor is controlled by the parameter $\epsilon$, the forcing, and the initial conditions. We are particularly interested in configurations where the linear system (52) is highly non-normal which may result in transient growth of the solutions of the system (52) and a significant model error of a scalar Mean Stochastic Model for the observed variable which is discussed in Section ???. It is important to understand the main properties of the system (52) in various dynamical regimes, since the short and intermediate range forecast skill of the MSM and the model error depend on these regimes.

There are two special limiting configurations:

- **Eigenvectors parallel** (Highly non-normal behaviour; transient growth possible)

  We first write the solution to (52) using Duhamel’s formula as

  $\displaystyle y(t) = e^{-b(t-t_0)}y_0 + \int_{t_0}^{t} e^{-b(t-s)}(\epsilon^{-1}x(s) + F_y(s))ds + \sigma \int_{t_0}^{t} e^{-b(t-s)}dW(s),$  

  (83)

  and consequently

  $\dot{x} = -a x(t) + F_x(t) + \int_{t_0}^{t} e^{-b(t-s)}x(s)ds$

  $+ \epsilon e^{-b(t-t_0)} \left( y_0 + \int_{t_0}^{t} e^{-b(s-t_0)} F_y(s)ds \right) + \epsilon \sigma \int_{t_0}^{t} e^{-b(t-s)}dW(s)$  

  (85)

  Clearly, when $\epsilon \sigma \to \infty$ a lot of variance is ‘pumped’ into the observed dynamics. In the limit $\epsilon \sigma \to 0$ the resolved dynamics is deterministic. It is useful to consider these two limits separately.

- $\epsilon \ll 1$ This configuration corresponds to non-normal dynamics with eigenvectors $v_{1,2}$ of $\hat{L}$ nearly parallel (see figure 1). As can be seen from (59)

  $\lim_{\epsilon \to 0} v_{1,2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$

  (86)

  while the eigenvalues (55) remain constant. The eigenvectors of the equilibrium covariance matrix (68) whose (orthogonal) eigenvectors become

  $\lim_{\epsilon \to 0} w_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \lim_{\epsilon \to 0} w_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$

  (87)

  while the limiting values of the respective eigenvalues are

  $\lim_{\epsilon \to 0} \xi_1 = 0, \quad \lim_{\epsilon \to 0} \xi_2 = \frac{\sigma^2(\lambda_1\lambda_2 + a^2)}{2(\lambda_1 + \lambda_2)\lambda_1\lambda_2}.$

  (88)

  If we consider the eigenvectors of the equilibrium covariance matrix (68) as the climatological basis, the dominant EOF for $\epsilon \ll 1/\sigma$ is given by $w_2$ which in the limit $\epsilon \to 0$ is aligned with the unresolved dynamics and coincides with the eigenvectors $v_{1,2} \sim (0, 1)^T$ of the system (52).
Figure 1: Example of mean (non-normal) dynamics of the two-dimensional system (52) for $\epsilon \ll 1$.

- $\epsilon \gg 1$ (See figure 2 for an example.) In this case the eigenvectors $v_{1,2}$ of $\hat{L}$ tend to

$$\lim_{\epsilon \to \infty} v_{1,2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

which is clear from (59). The eigenvectors of the equilibrium covariance matrix (68) become

$$\lim_{\epsilon \to \infty} w_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \lim_{\epsilon \to \infty} w_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

while the limiting values of the respective eigenvalues are

$$\lim_{\epsilon \to \infty} \xi_1 = \frac{\sigma^2}{2(\lambda_1 + \lambda_2)}, \quad \lim_{\epsilon \to \infty} \xi_2 = \infty.$$

If we consider the eigenvectors of the equilibrium covariance matrix (68) as the climatological basis, the dominant EOF for $\epsilon \gg 1$ is given by $w_2$ which in the limit $\epsilon \to \infty$ is aligned with the resolved dynamics and coincides with the eigenvectors of the system $v_{1,2} \sim (1,0)^T$ of the system (52). If $\epsilon \gg 1/\sigma$ a lot of variance is pumped from the unresolved modes into the resolved dynamics.

- **Eigenvectors orthogonal** (Normal behaviour, no transient growth; see figure 3) When $\epsilon = 1$, the
eigenvectors of the system matrix $\hat{L}$ are orthogonal and given by

$$v_1 = \left( \frac{1}{\lambda_1 + a} \right), \quad v_2 = \left( \frac{1}{\lambda_2 + a} \right). \tag{92}$$

It can be easily verified from (60) that the angle between the two vectors is $\alpha(\epsilon = 1) = \pi/2$. The climatological basis vectors are in this case given by

$$w_1 = \left( \frac{1}{\lambda_1 \lambda_2 + a^2 - 1 - \sqrt{(\lambda_1 \lambda_2 + a^2 - 1)^2 + 4a^2}} \right), \tag{93}$$

$$w_2 = \left( \frac{2a}{\lambda_1 \lambda_2 + a^2 - 1 + \sqrt{(\lambda_1 \lambda_2 + a^2 - 1)^2 + 4a^2}} \right), \quad w_1 \cdot w_2 = 0. \tag{94}$$

### 2.2 Different models with the same marginal statistics

We discuss here some consequences of taking account of only a subset of variables in a stochastic system which is a generic situation when constructing simplified models of complex high-dimensional systems on some coarse-grained subset of variables. Here, we illustrate the consequences of non-uniqueness associated in the simplest possible case, i.e. within the family of two-dimensional models discussed above. These issues will be crucial for many aspects of uncertainty quantification for more complex systems discussed in the following lectures.

First we provide the setting for our considerations:
Definition [Marginal probability density]: Suppose that $X : \Omega \to \mathbb{R}^n$ is a random variable with probability density function $p(x_1, \ldots, x_m)$ where $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$. The marginal probability density $\tilde{p}(x_i)$ for $x_i, i \in [1, \ldots, n]$ is given by

$$\tilde{p}(x_i) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} p(y_1, \ldots, x_i, \ldots, y_n) dy_1 \ldots dy_{i-1} dy_{i+1} \ldots dy_n. \quad (95)$$

Analogous definition holds for a marginal density for any subset of components of $x$.

Remark. Standard proofs show that for a Gaussian random variable with mean $\mu$ and covariance $\Sigma$ the marginal density over a subset of components is obtained by simply dropping the remaining components, e.g. given a Gaussian vector $(x_1, x_2, x_3)^T$ with $\mu = (\mu_1, \mu_2, \mu_3)^T$ and covariance

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix}, \quad (96)$$

the marginal density for $(x_1, x_3)$ is given by a Gaussian with

$$\tilde{\mu} = \begin{pmatrix} \mu_1 \\ \mu_3 \end{pmatrix}, \quad \tilde{\Sigma} = \begin{pmatrix} \Sigma_{11} & \Sigma_{13} \\ \Sigma_{31} & \Sigma_{33} \end{pmatrix}. \quad (97)$$

Example: Nonuniqueness of marginal equilibrium dynamics in the class of two-dimensional systems (52)
Consider now the marginal equilibrium dynamics for \( x(t) \) in the system (52). This example is particularly appropriate from the point of view of applications to uncertainty quantification in Atmosphere Ocean Science since in one usually has only available the statistics of the natural system on its attractor. Recall that for constant forcing the equilibrium statistics of (52) is given by (73), (74), (67), (72). Consequently, the marginal density for the equilibrium statistics of \( x(t) \) in (52) is given by a Gaussian with

\[
\begin{align*}
\mu_1 &= \langle x \rangle_{eq} = \frac{\lambda_1 b F_0^x}{\lambda_1 \lambda_2} + \frac{\epsilon F_0^y}{\lambda_1 \lambda_2}, \\
\Sigma_{11} &= -\frac{\sigma^2 \epsilon^2}{2(\lambda_1 + \lambda_2) \lambda_1 \lambda_2},
\end{align*}
\]  

(98)  

(99)

and the autocorrelation function for \( x(t) \) at equilibrium is given by

\[
(C_{att})_{11} = \frac{\lambda_1 e^{\lambda_2 \tau} - \lambda_2 e^{\lambda_1 \tau}}{\lambda_1 - \lambda_2}.
\]  

(100)

The following exercise is a prelude to UQ: Assume that the system (52) with parameters \( \{a, b, \epsilon, \sigma, F_0^x, F_0^y\} \) represents the truth and that we know the structure of the perfect model but not its parameters. Moreover, assume that the only information we can use to estimate these parameters is from the marginal equilibrium statistics for \( x(t) \). Thus we seek a model of the truth system with the structure (52) and parameters \( \{\tilde{a}, \tilde{b}, \tilde{\epsilon}, \tilde{\sigma}, \tilde{F}_0^x, \tilde{F}_0^y\} \) such that the equilibrium statistics of the imperfect model coincides with (98)-(100) which leads to the following constraints

\[
\begin{align*}
\tilde{\lambda}_{1,2} &= \lambda_{1,2}, \\
\tilde{b} F_0^x + \tilde{\epsilon} F_0^y &= b F_0^x + \epsilon F_0^y, \\
\tilde{\sigma}^2 \tilde{\epsilon}^2 &= \sigma^2 \epsilon^2.
\end{align*}
\]  

(101)  

(102)  

(103)

There are in total four constraints and six parameters in the considered model (52). Thus, we can conclude that there exists a two-parameter family of models (52) with the same marginal equilibrium statistics as the truth system. One particularly revealing choice of parameterization, with parameters \( (w_1, w_2) \), is

\[
\begin{align*}
\tilde{\sigma} &= w_1, \\
\tilde{\epsilon} &= \frac{\sigma \epsilon}{w_1}, \\
\tilde{\alpha} &= a, \\
\tilde{b} &= b, \\
\tilde{F}_0^x &= w_2, \\
\tilde{F}_0^y &= \frac{w_1}{\sigma \epsilon} \left(b (F_0^x - w_2) + \epsilon F_0^y\right).
\end{align*}
\]  

(104)  

(105)  

(106)  

(107)  

(108)  

(109)

Note, in particular, that for any fixed \( \epsilon \) in the true system we can choose \( \tilde{\epsilon} \) such that the imperfect model will display dynamics ranging from normal to very nonnormal. It is also worth stressing that the situation considered in this example corresponded to the case when the structure and dimensionality of the imperfect models coincided with those for the truth system. Later in the course we will discuss more realistic situations when the imperfect models only take into account a subset of dynamical variables of the truth system. Again the two-dimensional prototype model discussed in this section will serve as a very useful testbed for illustrating various effects of dimensionality reduction in the simplest possible setting.
3 Non-Gaussian systems

In the previous section we thoroughly discussed an example of a linear two-dimensional stochastic system with Gaussian dynamics. Recall that the probability measure on the attractor (provided the attractor exists) for any stochastic process satisfying a linear SDE with deterministic coefficients and additive noise is necessarily Gaussian. Systems with Gaussian equilibrium measure and Gaussian initial conditions (i.e. drawn from the normal distribution) are referred to as Gaussian systems. Here, we consider a qualitatively different dynamics of a real scalar \( u(t) \) given by a linear SDE with multiplicative and additive noise. In particular, we show that the associated dynamics is non-Gaussian for any non-zero amplitude of the multiplicative noise regardless of the nature of the mean dynamics (i.e. linear or nonlinear) and even for Gaussian initial conditions. Besides their educational value as unambiguous models illustrating the effects of non-Gaussianity, such scalar models often represent limiting scenarios in much more complex nonlinear systems when the dynamics at some scales approaches the white noise limit. We formally derive here:

(i) Equilibrium PDFs for two models (one linear and one nonlinear) with multiplicative noise in the case when the mean damping and forcing are constant,

(ii) The moments of the time-dependent PDF and condition for the existence of all moments of order up to \( n \).

The main focus of the following discussion is on the connection between intermittency of the dynamics and fat-tailed PDFs with small number of finite moments.

3.1 Linear scalar model with multiplicative noise

We consider here the following scalar linear SDE with multiplicative and additive noise

\[
du(t) = [-au(t) + f]dt + bu(t)dW_b(t) + c dW_c,
\]  

(110)

where \( W_b \) and \( W_c \) are independent scalar Wiener processes. The coefficients \( a, b, c \) and the forcing \( f \) are, in general, time-dependent deterministic functions but we restrict the analysis below to the time-independent case.

3.2 Exact path-wise solutions for the system (110)

The system (110) can be solved exactly. Consider first the homogeneous problem associated with (110) given by

\[
du(t) = -au(t)dt + bu(t)dW_b(t),
\]  

(111)

with \( a, b \) known deterministic functions of time. The homogeneous equation (111) can be solved by substituting, \( y = \ln u \), so that, with the help of the usual Ito rules, equation (111) becomes

\[
dy = \frac{du}{u} - \frac{1}{2} \frac{(du)^2}{u^2} = -(a + \frac{1}{2} b^2)dt + b dW_b(t).
\]  

(112)

The above equation is directly integrable and, consequently, equation (111) has exact solutions in the form

\[
u(t) = u_0 \exp \left(- \int_{t_0}^t (a(s) + \frac{1}{2} b(s)^2)ds + \int_{t_0}^t b(s)dW_c(s) \right) = u_0 \Phi(t),
\]  

(113)

where \( u_0 \equiv u(t_0) \). Consider now the inhomogeneous equation (110) and write
\[ z(t) = u(t) \Phi^{-1}(t), \]

with \( \Phi(t) \) defined in (113). Noting that \( d(\Phi^{-1}) = -(d\Phi)\Phi^{-2} + (d\Phi)^2\Phi^{-3} \), we use again Ito calculus to evaluate

\[
dz = du \Phi^{-1} + u d(\Phi^{-1}) + du d(\Phi^{-1}),
\]

which leads to a directly integrable equation

\[
dz = \frac{1}{\Phi(t)} (f(t) dt + c dW_c).
\]

Hence,

\[
u(t) = \Phi(t) \left( u_0 + \int_{t_0}^{t} \frac{f(s)}{\Phi(s)} ds + \int_{t_0}^{t} dW_c(s) \right).\]

### 3.3 Invariant measures for the system (110)

The invariant measure associated with (110) is given by the time-independent solutions of the Fokker-Planck equation

\[
\partial_t p(u,t) = -\partial_u [(au + f)p(u,t)] + \frac{1}{2} \partial_u^2 \left[(b^2 u^2 + c^2) p(u,t) \right] \equiv L_{FP} p(u,t).
\]

The stationary solutions of (117), i.e. \( L_{FP} p_{eq} = 0 \), satisfy

\[
\partial_u \left[(b^2 u^2 + c^2) p_{eq}(u) \right] = 2(-au + f) p_{eq}(u),
\]

which can be easily integrated to yield

\[
p_{eq}(u) = \frac{N_0}{A(u)} \exp \left( \int_{u_0}^{u} B(u')/A(u') du' \right)
\]

where

\[
A(u) = bu^2 + c^2, \quad B(u) = 2(-au + f).
\]

There exist three distinct classes of invariant measures associated with (119) depending on the parameters \( a, b, c \):

(I) \( a > 0, b \neq 0, c \neq 0 \) (Non-zero additive and multiplicative noise; non-Gaussian PDF with fat algebraic tails.) In this case the solution of (119) yields a PDF with fat algebraic tails and is given by

\[
p_{eq}(u) = \frac{N_0}{(b^2 u^2 + c^2)^{1+\alpha/2}} \exp \left( \frac{2f}{|bc|} \arctan \left( \frac{|b|}{|c|} u \right) \right),
\]

with \( N_0 \) a normalizing constant; see figures 5 and 6 for examples.

It can be easily verified that for \( b \neq 0, c = 0 \) the invariant measure does not exist. The solution of (119)

\[
p_{eq}(u) = \frac{N_0}{u^{2(1+\alpha/b^2)}} \exp \left( -2f/(b^2 u) \right)
\]

does not exist at \( u = 0. \)
Figure 4: a) Exact equilibrium PDF (123) (solid black), its numerical approximation (red) and a Gaussian approximation (with matching mean and variance; dashed) for the system (110) in a Gaussian regime with no multiplicative noise, b) typical path-wise solution.

System parameters: $a = 5, b = 0, c = 0.5, f = 0.6$,
Figure 5: a) Exact equilibrium PDF (121) (solid black), its numerical approximation (red) and a Gaussian approximation (with matching mean and variance; dashed) for the system (110) in a strongly non-Gaussian regime with multiplicative and additive noise and non-zero forcing, b) typical path-wise solution.

System parameters: $a = 5, b = 1, c = 0.3, f = 0.6,$
Figure 6: a) Exact equilibrium PDF (121) (solid black), its numerical approximation (red) and a Gaussian approximation (with matching mean and variance; dashed) for the system (110) in a strongly non-Gaussian regime with multiplicative and additive noise and zero forcing, b) typical path-wise solution.

System parameters: $a = 5, b = 1, c = 0.5, f = 0,$
(II) $a > 0, b = 0, c \neq 0$ (Only additive noise; Gaussian PDF.) In this case the solution of (119) yields a Gaussian density

$$p_{eq}(u) = N_0 \exp \left( -\frac{a}{c^2} (u - f/a)^2 \right).$$

(123)

An example of such density is shown in figure 4. Note also that for $a = 0$ the invariant measure does not exist.

(III) $a > 0, b = 0, c = 0$ (Deterministic case; no noise.) In this trivial case

$$p_{eq}(u) = \delta(u - f/a),$$

(124)

where $\delta(\cdot)$ is the Dirac delta function.

3.4 Approach to the equilibrium PDF

The invariant measures derived in the previous section describe the limiting dynamics of the system (110). Clearly, for arbitrary statistical initial conditions (i.e. the initial density $p(t = t_0)$) the system will evolve towards the statistical attractor (for constant forcing $f$ the attractor for the mean dynamics is a fixed point and the equilibrium statistics is time-independent).

In figures 7-8 we show two examples of such evolution towards the equilibrium for two distinct dynamical configurations. The first example shown in figure 7 corresponds to the evolution of (110) from Gaussian initial conditions when $b \ll c$ and $f = 0$ so that the invariant measure is nearly Gaussian. The numerical approximations are obtained through Monte Carlo simulations rather than solving the Fokker-Planck equation.

The evolution illustrated in figure 8 corresponds to the evolution of (110) from Gaussian initial conditions when $b \gg c$ and $f \neq 0$ so that the invariant measure is non-Gaussian and skewed. The numerical approximations are obtained through Monte Carlo simulations.
ps(ℜe[Tk])

\( t = 0.2 \)
\( \Re[\text{Tk}] \)
\( \eta_U = 3 \), \( \sigma_{Tk} = 1 \)

\( t = 0.5 \)
\( \Re[\text{Tk}] \)

\( t = 1 \)
\( \Re[\text{Tk}] \)

\( t = 2 \)
\( \Re[\text{Tk}] \)

\( t = 3 \)
\( \Re[\text{Tk}] \)

Figure 7: Time-dependent evolution of the PDF associated with the unforced linear system (110) with multiplicative noise in the nearly-Gaussian regime (large damping \( a \) and small multiplicative noise \( b \ll c \)). The initial PDF is Gaussian and the panels show snapshot at different stages of transient onto the equilibrium PDF. The PDFs were estimated numerically based on 40000 sample paths with Gaussian initial conditions.

Figure 8: Similar to above but this time there is a significant multiplicative noise, \( b \gg c \), in the system (110) and nonzero forcing \( f \neq 0 \) leading to a non-Gaussian skewed equilibrium PDF.
3.5 Moments of the time-dependent PDFs associated with (110).

We assume here that \( b = \text{const}, c = \text{const} \) in (110) and the forcing is bounded, i.e. \( \max_{t \in [t_0, \infty)} f(t) < \infty \). We show that for time-periodic damping \( a(t) \) with period \( T \) and a finite positive mean

\[
0 < a_0 \equiv \frac{1}{T} \int_{t_0}^{t_0+T} a(s)ds,
\]

the moments of the time-dependent PDFs of the system (110) exist up to order \( n \) provided that \( c \neq 0 \) and either \( b \neq 0 \) or \( a \neq 0 \) and

\[
-na_0 + \frac{1}{2}n(n-1)b^2(t-t_0) < 0.
\]

The \( n \)-th moment of (110) is given by

\[
\langle u^n(t) \rangle = \langle u^n_0 \rangle e^{J_n(t_0,t)} + \int_{t_0}^{t} e^{J_n(s,t)} \left( nf(s)\langle u^{n-1}(s) \rangle + \frac{1}{2}n(n-1)c^2\langle u^{n-2}(s) \rangle \right) ds,
\]

where

\[
J_n(s,t) = -n \int_{s}^{t} a(t')dt' + \frac{1}{2}n(n-1)b^2(t-s).
\]

In the above formulas we additionally set \( \langle u(t)^{n-k} \rangle = 1 \) for \( n = k, k = 1,2 \) and \( \langle u(t)^{n-k} \rangle = 0 \) for \( n - k < 0, k = 1,2 \). Provided that (125) holds, the moments of the PDF on the statistical attractor are given by

\[
\langle u^n(t) \rangle_{\text{att}} = \lim_{t_0 \to -\infty} \int_{t_0}^{t} e^{J_n(s,t)} \left( nf(s)\langle u^{n-1}(s) \rangle + \frac{1}{2}n(n-1)c^2\langle u^{n-2}(s) \rangle \right) ds,
\]

We find the moments of the probability distribution associated with this scalar system with multiplicative noise directly from the dynamical equation (110) and the Ito formula. Note first that, using Ito rules, we have

\[
du^n = nu^{n-1}du + \frac{n(n-1)}{2}u^{n-2}(du)^2
\]

which upon evaluation on (110) and taking the ensemble average leads to

\[
\langle u^n(t) \rangle = \langle u^n_0 \rangle e^{J_n(t_0,t)} + \int_{t_0}^{t} e^{J_n(s,t)} \left( nf(s)\langle u^{n-1}(s) \rangle + \frac{1}{2}n(n-1)c^2\langle u^{n-2}(s) \rangle \right) ds
\]

where

\[
J_n(s,t) = -n \int_{s}^{t} a(t')dt' + \frac{1}{2}n(n-1)b^2(t-s).
\]

In the above formulas we additionally set \( \langle u(t)^{n-k} \rangle = 1 \) for \( n = k, k = 1,2 \) and \( \langle u(t)^{n-k} \rangle = 0 \) for \( n - k < 0, k = 1,2 \). For time periodic damping \( a(t) = a_0 + \bar{a}(t) \) with \( \bar{a}(t) \) having zero time average, (131) can be written as

\[
J_n(s,t) = (-na_0 + \frac{1}{2}n(n-1)b^2) (t-s) - n \int_{s}^{t} \bar{a}(t')dt.
\]

The last term in (132) is bounded since it is a periodic function with zero mean. Consequently, provided that

\[
-na_0 + \frac{1}{2}n(n-1)b^2 < 0,
\]

the first term in (130) is finite for any \( t \in [t_0, \infty) \).
Consider now the \( n \)-th moment and assume that \( f(t), \langle u^{n-1}(t) \rangle, \langle u^{n-2}(t) \rangle \) exist and
\[
\max_{t \in [t_0, \infty)} \langle u^{n-1}(t) \rangle < C_1, \quad \max_{t \in [t_0, \infty)} \langle u^{n-2}(t) \rangle < C_2, \quad \max_{t \in [t_0, \infty)} f(t) < C_f
\] (134)
Then, provided that (133) holds, we have the following estimate on the second term in (130)
\[
\int_{t_0}^{t} e^{J_n(s,t)} (n f(s) \langle u^{n-1}(s) \rangle + \frac{1}{2} n(n - 1) c^2 \langle u^{n-2}(s) \rangle) \, ds \leq \\
\leq (n C_f C_1 + \frac{1}{2} n(n - 1) c^2 C_2) \int_{t_0}^{t} e^{-n \int_{s}^{t} \tilde{a}(t') dt'} \, ds < \infty. \quad (135)
\]
The last estimate in (135) arises from the fact that \( \int_{s}^{t} \tilde{a}(t') dt' \) is a bounded function for \( \tilde{a}(t) \) periodic with zero time average.

Finally, it is clear that provided that (125) holds, the moments of the PDF on the statistical attractor are given by
\[
\langle u^n(t) \rangle_{\text{att}} = \lim_{t_0 \to -\infty} \int_{t_0}^{t} e^{J_n(s,t)} (n f(s) \langle u^{n-1}(s) \rangle + \frac{1}{2} n(n - 1) c^2 \langle u^{n-2}(s) \rangle) \, ds. \quad (136)
\]
In figures 9-10 we illustrate the possible rich behaviour of the time-dependent PDFs on the attractor of (110) for time-periodic damping \( a \) and forcing \( f \). These figures show the first four normalized moments of the time-dependent PDFs given by
- Mean (on attractor)
  \[ \mu = \langle u(t) \rangle, \] (137)
- Standard deviation (on attractor)
  \[ \sigma = \sqrt{\langle (u(t) - \mu(t))^2 \rangle}, \] (138)
- Skewness (on attractor)
  \[ \gamma = \frac{\langle (u(t) - \mu)^3 \rangle}{\sigma^3} = \frac{\langle u^3(t) \rangle - 3 \mu \langle u^2(t) \rangle + 2 \mu^3}{\sigma^3}, \] (139)
- Excess Kurtosis (on attractor)
  \[ \gamma_2 = \frac{\langle (u(t) - \mu)^4 \rangle}{\sigma^4} - 3 = \frac{\langle u^4(t) \rangle - 4 \mu \langle u^3(t) \rangle + 6 \mu^2 \langle u^2(t) \rangle - 3 \mu^4}{\sigma^4} - 3. \] (140)
The moments about the origin, \( \langle u^n(t) \rangle \), are given by (136). Note that the corresponding PDFs on the attractor are time-dependent and non-Gaussian which is highlighted by a non-zero time-dependent skewness and large kurtosis in figures 9-10.
3.5.1 Moments of the equilibrium PDFs for constant mean damping and forcing

When \( a \) and \( f \) in (110) are constant, the first three moments of the PDF associated with this system are

\[
\langle u(t) \rangle = \langle u_0 \rangle + \frac{f}{J_1} e^{J_1(t-t_0)} - \frac{f}{J_1},
\]

\( (141) \)

\[
\langle u^2(t) \rangle = e^{J_2(t-t_0)} \left( \langle u_0^2 \rangle + \frac{c^2}{J_2} - \frac{2f(J_2(u_0) + f)}{J_2(J_1 - J_2)} \right) + e^{J_1(t-t_0)} \frac{2f(J_1(u_0) + f)}{J_1(J_1 - J_2)} + \frac{2f^2 - J_1 c^2}{J_1 J_2},
\]

\( (142) \)

\[
\langle u^3(t) \rangle = e^{J_3(t-t_0)} \times \left( \langle u_0^3 \rangle + \langle u_0^2 \rangle \frac{3f}{J_2 - J_3} + \langle u_0 \rangle \frac{6(2f^2 - (J_2 - J_3) c^2)}{(J_2 - J_3)(J_1 - J_3)} + \frac{3f(c^2(2J_3 - J_1 - J_2)) + 2f^2}{J_3(J_2 - J_3)(J_1 - J_3)} \right)
\]

\[
+ e^{J_2(t-t_0)} \times \frac{3f}{J_2 - J_3} \left( \langle u_0^3 \rangle + \frac{c^2}{J_2} - \frac{2f}{J_1 - J_2} \left( \langle u_0 \rangle + \frac{f}{J_2} \right) \right)
\]

\[
+ e^{J_1(t-t_0)} \times \frac{3(c^2(J_1 - J_2) + 2f^2)}{(J_1 - J_2)(J_1 - J_3)} \left( \langle u_0 \rangle + \frac{f}{J_1} \right)
\]

\[
+ \frac{3f((J_1 + J_2) c^2 - 2f^2)}{J_1 J_2 J_3},
\]

\( (143) \)

where \( \langle u_0 \rangle \) denotes the initial statistical condition.

The moments of the invariant measure, can be obtained in the limit \( t_0 \to -\infty \), and provided that

\[
\chi_n = -n a_0 + \frac{1}{2} n(n - 1) b^2 < 0,
\]

\( (144) \)

for \( n \leq 4 \), the first four moments exist and are given by

\[
\langle u \rangle = - \frac{f}{\chi_1},
\]

\( (145) \)

\[
\langle u^2 \rangle = \frac{2f^2 - \chi_1 c^2}{\chi_1 \chi_2},
\]

\( (146) \)

\[
\langle u^3 \rangle = \frac{3f((\chi_1 + \chi_2) c^2 - 2f^2)}{\chi_1 \chi_2 \chi_3},
\]

\( (147) \)

\[
\langle u^4 \rangle = \frac{6 c^4}{\chi_2 \chi_4} + \frac{12f^2(2f^2 - c^2(\chi_1 + \chi_2 + \chi_3))}{\chi_1 \chi_2 \chi_3 \chi_4},
\]

\( (148) \)
Figure 9: The first four normalized moments (137)-(140) for the scalar model (110) with non-zero multiplicative noise when both the damping $a$ ($\hat{\gamma}$ in the legends) and the forcing $f$ are deterministic and time-periodic.

Figure 10: Similar to the above but here the damping $a$ ($\hat{\gamma}$ in the legends) is much smaller and the system and the time-dependent equilibrium measure on the attractor is very non-Gaussian and intermittent (see the large values of the kurtosis).
3.6 Canonical scalar model with cubic nonlinearity and deterministic structural instability

3.6.1 Motivation

We introduce here a nonlinear generalization of the scalar model with multiplicative noise discussed in §3.1. This model is a normal form for a reduced stochastic climate model for the low-frequency variability, as discussed in [8] and can be used for studying low-frequency patterns such as the North Atlantic Oscillation. Such normal forms can be obtained via the so-called stochastic mode reduction in the case when the dyad interactions along with the triad interactions between the low- and high-frequency subspaces of the turbulent dynamics are considered. This procedure results in a correlated additive and multiplicative (CAM) noise terms and strong nonlinear dissipation. The rich dynamical structure of the model (CAM and strong nonlinearity) including deterministic structural instability allows one to consider various dynamical regimes, from nearly Gaussian to very non-Gaussian, and it provides a good toy model for studying the Fluctuation Dissipation Theorem (FDT; this is a topic for future lectures). Moreover, in the strongly non-Gaussian regimes, with nonzero skewness and high kurtosis, are plausible systems for climate modeling.

3.6.2 The model and its statistics

As discussed [8], the simplest reduced stochastic model for a scalar low-frequency climate variable obtained via the so-called stochastic mode reduction is given by

\[ dx = [F + ax + bx^2 - cx^3]dt + (A - Bx)dW + \sigma dW_A \]  

where \( W \) and \( W_A \) are independent scalar Wiener processes; the term \((A - Bx)\dot{W}\) is referred to as the CAM noise (Correlated Additive and Multiplicative noise) since its presence, as well as the cubic terms, are associated with the same physical process. Following [8], we will refer to this system as the normal form for scalar stochastic climate models. The corresponding Fokker-Planck equation for the evolution of the probability density of the process \( x \) solving (149) becomes

\[ \frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}[(F + ax + bx^2 - cx^3)p] + \frac{1}{2} \frac{\partial^2}{\partial x^2}[(A - Bx)^2 + \sigma^2]p. \]  

For the case of nonzero CAM noise, we find the following equilibrium PDF (see [8] for details)

\[ p_{eq}(x) = \frac{N_0}{((Bx - A)^2 + \sigma^2)a_1} \exp(d \arctan \left( \frac{Bx - A}{\sigma} \right)) \exp(-\frac{c_1 x^2 + b_1 x}{B^4}), \]  

where \( N_0 \) is a normalizing constant, and

\[ a_1 = 1 - \frac{-3A^2c + AB^2 + 2AB + c\sigma^2}{B^4}, \]  

\[ b_1 = 2B^2 - 4cAB, \]  

\[ c_1 = cB^2, \]  

\[ d = \frac{d_1}{\sigma} + d_2\sigma, \]  

\[ d_1 = 2\frac{A^2B - A^3c + AAB^2 + \sigma}{B^4}, \]  

\[ d_2 = \frac{6cA - 2bB}{B^4}. \]
On the other hand, in a special case of the additive noise only, i.e., \( A = B = 0 \), we find the following invariant PDF

\[
p_{eq}(x) = N_0 \exp \left( \frac{2}{\sigma^2} \left( Fx + \frac{a}{2}x^2 + \frac{b}{3}x^3 - \frac{c}{4}x^4 \right) \right)
\]  

(158)

### 3.6.3 Regime study

Here, we employ the classical ODE stability analysis to the deterministic part of (149) in order to find the fixed points for different values of the parameters \( F, a, b, \) and \( c \). It turns out that for fixed values of \( b \) and \( c \), the two-dimensional parameter space \( a - F \) divides into two parts. For the parameters that belong to the first part of the space \( a - F \), the deterministic part \( F + ax + bx^2 - cx^3 \) has three real roots, implying that the system

\[
\frac{dx}{dt} = F + ax + bx^2 - cx^3
\]  

(159)

has three fixed points, two stable and one unstable (for \( c < 0 \)). On the other hand, the second part of the set consists of the parameters, for which the system only has one stable fixed point (again, for \( c < 0 \)). On the border of these two subspaces of the \( a - F \) space, there are parameters, for which two of the three fixed points coincide and we have one stable fixed point and one unstable (stable from one side and unstable from another). The boundary between these two different regimes in the parameter space is given by a singular fold set for which there are exactly two fixed points. It turns out that here, we can analytically obtain the singular fold set. First, we write the cubic polynomial in a canonical form (for
$c \neq 0$)

$$x^3 + a_2 x^2 + a_1 x + a_0 = 0,$$  \hspace{1cm} (160)

where $a_2 = -b/c$, $a_1 = -a/c$, $a_0 = -F/c$. From the theory for the cubic polynomials it is known, that this polynomial has two real roots if and only if the determinant of this equation is zero. The determinant has the form

$$D = \left( \frac{p}{3} \right)^3 + \left( \frac{q}{3} \right)^2$$  \hspace{1cm} (161)

where

$$p = \frac{1}{3}(3a_1 - a_2), \quad q = \frac{1}{27}(9a_1a_2 - 27a_0 - 2a_3^2).$$  \hspace{1cm} (162)

After substituting the parameters (160), we find that $D = 0$ iff

$$F = -\frac{ab}{3c} - \frac{2b^3}{27c^2} \pm \left( \frac{a}{3c} + \frac{b^2}{9c^2} \right)^{3/2}.$$  \hspace{1cm} (163)

We note that $D = 0$ is possible only for

$$a > a_c \equiv -\frac{b^2}{3c}.$$  \hspace{1cm} (164)

for positive $c$. Then, for any $a > a_c$, we find two values of $F_{1,2}$ corresponding to the fixed points on the fold set. Assume that $F_1 \leq F_2$. Then, for all $F_1 < F < F_2$ the system has three fixed points. Otherwise, there is only one fixed point, since the corresponding cubic equation has one real and two complex conjugate roots (see figure 11 for examples).

We will return to this model in the subsequent lectures.

## 4 Building complex models with regime switching

We present here a useful procedure for a systematic building of a hierarchy of models with increasing complexity which are capable of mimicking realistic dynamics while retaining analytical tractability. The dynamics of the models discussed below are characterized by a stochastic large-scale regime switching with small-scale noise and are based on the Langevin-type equation (i.e. linear SDE with Gaussian solutions) in which some parameters are given by the so-called finite-state Markov jump process which is introduced below. A widely studied example of such a regime-like behaviour in the atmosphere and ocean science is associated with monitoring one Rossby wave of a baroclinic fluid as the mean state randomly changes between baroclinic stability and instability [10].

We apply this procedure in two simple cases related to stochastic models of turbulence with regime switching. The first simpler example corresponds to the dynamics of a single Fourier mode in a turbulent signal, while the second example shows some basic properties of the full PDE case of the spatially extended system.

### 4.1 Markov jump processes

We consider here a subclass of Markov jump process with a finite state set, which are given by a continuous time stochastic process taking at any time $t \in \mathbb{R}$ one of the discrete values in the finite set of states

$$S = \{s_i\}_{i \in \mathbb{N}}, \quad \mathbb{N} = \{1, 2, \ldots, N\}, \quad 0 < N < \infty.$$  \hspace{1cm} (165)

We assume that the set $S$ is nontrivial in the sense that at least one of the states $s_i$ is different from the others.
Figure 12: Sample path of a scalar two-state Markov jump process with one state positive, one state negative, and a positive mean (182). The sample paths of this process are discontinuous; the vertical lines connecting the positive and negative states are for visualization purposes only. In this example the switching rate $\mu$ from the negative to the positive state is greater than the switching rate for the complementary event (i.e. transition from the positive to the negative state.)

In contrast to the diffusion-type stochastic processes considered earlier (see §1.2), non-trivial Markov jump processes have discontinuous paths and do not satisfy the Fokker-Planck equation. Instead, the probability densities associated with such processes satisfy the so-called Master equation which is derived from the differential version of the Kolmogorov equation (see e.g. [3] where it is referred to as Chapman-Kolmogorov equation) upon the requirement that the process transitions between isolated states (the spectrum of the states can be continuous though for a general Markov jump process). Below we will obtain the Master equation for the simplest possible case of a two-state Markov jump process which is particularly interesting because of its exact solvability (which is uncommon).

Consider now the simplest case of a Markov jump process $X(t)$ which, at any time $t \in \mathbb{R}$, takes one of only two values, i.e., the set of states is in this case

$$S = \{s_{st}, s_{un}\},$$

(166)

where, for simplicity, we assume that $s_{st}, s_{un} \in \mathbb{R}$ and call these states, respectively, the stable and unstable state. A sample path of such a process is shown in figure 12. Let $\nu$ be the rate of change from stable state $s_{st}$ to unstable state $s_{un}$. Similarly, we denote by $\mu$ the rate of change from the unstable state to the stable state. If, in addition to the Markov property of the process, i.e.

$$P(X(t) = s_i | X(\tau) = s_j) = P(X(t - \tau) = y | X_0 = x), \quad \forall \ t > \tau > 0.$$  

(167)

Thus, the process $X$ is fully determined by the transition probabilities $P(X(t) = s_i | X_0 = s_j)$, $i, j = 1, 2$. We note that $\nu$ and $\mu$ define the following local transition probabilities

$$P_{\Delta t}(s_{st}, s_{un}) := P(X(\Delta t) = s_{un} | X_0 = s_{st}) = \nu \Delta t + o(\Delta t),$$

(168)

$$P_{\Delta t}(s_{un}, s_{st}) := P(X(\Delta t) = s_{st} | X_0 = s_{un}) = \mu \Delta t + o(\Delta t),$$

(169)

$$P_{\Delta t}(s_{st}, s_{st}) := P(X(\Delta t) = s_{st} | X_0 = s_{st}) = 1 - \nu \Delta t + o(\Delta t),$$

(170)

$$P_{\Delta t}(s_{un}, s_{un}) := P(X(\Delta t) = s_{un} | X_0 = s_{un}) = 1 - \mu \Delta t + o(\Delta t).$$

(171)

Using the relations (168)-(171) and the assumed Markov property, the probability of ‘finding’ the process in the unstable state at $t + \Delta t$ is

$$P_{t + \Delta t}(s_{st}, s_{un}) = P_t(s_{st}, s_{un})P_{\Delta t}(s_{un}, s_{un}) + P_t(s_{st}, s_{st})P_{\Delta t}(s_{st}, s_{un})$$

$$= P_t(s_{st}, s_{un})[1 - \mu \Delta t] + P_t(s_{st}, s_{st})\nu \Delta t + o(\Delta t),$$

(172)

and, consequently, in the limit $\Delta t \to 0$ with the assumption that $P$ is at least $C^1$ w.r.t to $t$, one obtains the following differential equation (we skip a formal proof here)

$$\frac{\partial}{\partial t} P_t(s_{st}, s_{un}) = -\mu P_t(s_{st}, s_{un}) + \nu P_t(s_{st}, s_{st}).$$

(173)
The outcome of an analogous procedure for the remaining transition probabilities can be cast in the following form

\[
\frac{\partial}{\partial t} \hat{P}_t = \hat{P}_t A, \quad (174)
\]

\[
\hat{P}_t|_{t=0} = I, \quad (175)
\]

where \( I \) is a \( 2 \times 2 \) identity matrix and

\[
\hat{P}_t = \begin{pmatrix}
P_t(s_{st}, s_{st}) & P_t(s_{st}, s_{un}) \\
P_t(s_{un}, s_{st}) & P_t(s_{un}, s_{un})
\end{pmatrix}, \quad A = \begin{pmatrix}
-\nu & \nu \\
\mu & -\mu
\end{pmatrix}. \quad (176)
\]

Given that (174), is a linear system with constant coefficients, the probabilities for the system to be in stable and unstable regimes are easily found as

\[
P_t(s_{st}, s_{st}) = \frac{\mu}{\nu + \mu} + \frac{\nu}{\nu + \mu} e^{-(\nu + \mu)t}, \quad (177)
\]

\[
P_t(s_{st}, s_{un}) = \frac{\mu}{\nu + \mu} \left(1 - e^{-(\nu + \mu)t}\right), \quad (178)
\]

\[
P_t(s_{un}, s_{st}) = \frac{\nu}{\nu + \mu} \left(1 - e^{-(\nu + \mu)t}\right), \quad (179)
\]

\[
P_t(s_{un}, s_{un}) = \frac{\nu}{\nu + \mu} + \frac{\mu}{\nu + \mu} e^{-(\nu + \mu)t}, \quad (180)
\]

so that the two-state Markov jump process \( X(t) \) approaches a stationary process with probabilities

\[
P_{eq}(s_{st}, s_{st}) = \frac{\mu}{\nu + \mu}, \quad P_{eq}(s_{st}, s_{un}) = P_{eq}(s_{un}, s_{st}) = P_{eq}(s_{un}, s_{un}) = \frac{\nu}{\nu + \mu}. \quad (181)
\]

Note here that, unsurprisingly, the equilibrium probabilities do not depend on the initial states. Based on (181), the expectation for \( X(t) \) (see §1.1) is

\[
\bar{d} = E[X] = \frac{\mu s_{st} + \nu s_{un}}{\nu + \mu}. \quad (182)
\]

### 4.1.1 Switching times

We already know that, by definition, the two-state Markov jump process resides in one of the two states \( s_{st} \) or \( s_{un} \) and the switching rates between these states are \( \nu \) (for the \( s_{st} \to s_{un} \) transition) and \( \mu \) (for the \( s_{st} \to s_{un} \) transition). What can we say about the times when the process switches the state?

Suppose that the system transitioned to the stable state at some time \( t = t_0 \). We can define the time during which the system stays in the stable state as

\[
T_{st} = t^* - t_0, \quad t^* = \inf_{t \in [t_0, \infty)} \{t : X(t) = s_{un}\}. \quad (183)
\]

We will call \( t^* \) the switching time from the stable to unstable state and \( T_{st} \) the residence time in the stable state; clearly both \( t^* \) and \( T_{st} \) are random variables.

Note first that the probability that the process is still in \( s_{st} \) at time \( t = t_0 + \Delta t \) is

\[
P(T_{st} > t) = P(t^* > t_0 + t) = P(X(t_0 + t) = s_{st}, X(t_0) = s_{st}). \quad (184)
\]
Then, we can write
\[
P(X(t_0 + t + \Delta t) = s_{st}|X(t_0) = s_{st}) = \\
= P(X(t_0 + t) = s_{st}|X(t_0) = s_{st})P(X(t_0 + t + \Delta t) = s_{st}|X(t_0 + t) = s_{st}) \\
= P(X(t) = s_{st})P_{\Delta t}(s_{st}, s_{st}) = P(X(t_0 + t) = s_{st}|X(t_0) = s_{st})(1 - \nu\Delta t + o(\Delta t)),
\]
where the first equality is due to the assumed Markov property of the process, the second equality is due to the assumed homogeneity (167) and the last equality follows from (170). In the limit $\Delta t \to 0$ we obtain the following Master equation
\[
\frac{d}{dt}P(X(t_0 + t) = s_{st}|X(t_0) = s_{st}) = -\nu P(X(t_0 + t) = s_{st}|X(t_0) = s_{st}),
\]
so that
\[
P(X(t_0 + t) = s_{st}|X(t_0) = s_{st}) = e^{-\nu t}, \quad P(X(t_0 + t) = s_{un}|X(t_0) = s_{st}) = 1 - e^{-\nu t},
\]
which implies exponential distribution of $T_{st}$ with mean $1/\nu$
\[
P(T_{st} < t) = 1 - e^{-\nu t}.
\]
Similarly, we find that the time $T_{un}$ the system spends in the unstable regime before switching to the stable one is a random variable with the exponential distribution function
\[
P(T_{un} < t) = 1 - e^{-\mu t}.
\]

4.2 Two models with regime switching behaviour

4.2.1 Linear complex scalar dynamics with transient instabilites

Following [5], consider now an SDE for a complex scalar $u(t) \in \mathbb{C}$
\[
du = \left[ -\gamma(t)u + i\omega u + f(t) \right]dt + \sigma dW(t),
\]
where $W(t)$ is complex Wiener process with independent real and imaginary parts, and $f(t)$ is a known deterministic forcing. Here we regard $u(t)$ as one of the modes from a realistic turbulent signal as is often done in turbulence models [6, 2, 11, 7]. In order to mimic the intermittent instability which often occurs in nature, we allow $\gamma$ to switch between stable ($\gamma > 0$) and unstable ($\gamma < 0$) regimes according to a two-state Markov process. Such a switching process can mimic physical features of multiscale systems such as baroclinic instability [10].

Consider now the following hierarchy of models with increasing complexity:

(I) Constant damping $\gamma > 0$. This situation leads to familiar configuration of a linear SDE with path-wise solutions
\[
u(t) = e^{(-\gamma + i\omega)(t-t_0)}u_0 + \int_{t_0}^{t} f(s)e^{(-\gamma + i\omega)(t-s)}ds + \sigma \int_{t_0}^{t} e^{(-\gamma + i\omega)(t-s)}dW(s),
\]
The resulting complex scalar process $u(t)$ is clearly Gaussian as in the examples discussed in §1.2.3 and §2.
Figure 13: Sample paths of the non-Gaussian complex scalar process $u(t)$ satisfying (190) and of the damping $\gamma$ represented by scalar the two-state Markov jump process. Notice the transient instabilities in the dynamics of $u(t)$ and the corresponding bursts of energy (bottom panel).

(II) Deterministic time-dependent damping $\gamma$ such that $\int \gamma(s)ds < \infty$. This case is a generalization of (I) with path-wise solutions

$$u(t) = e^{-\int_{t_0}^{t} \gamma(t')dt' + i\omega(t-t_0)}u_0 + \int_{t_0}^{t} f(s)e^{-\int_{s}^{t} \gamma(t')dt' + i\omega(t-s)}ds + \sigma \int_{t_0}^{t} e^{-\int_{s}^{t} \gamma(t')dt' + i\omega(t-s)}dW(s),$$

(192)

Note that here might experience a significant growth due to the possibility of $\gamma < 0$. The process $u(t)$ remains Gaussian which is obvious from the path-wise solution (192) and the properties of the Ito stochastic integral.

(II) Damping with a stochastic component.

We highlight two different cases here:

- $\gamma(t)$ is given by some diffusion process and the path-wise solutions can be found in the same way as before (192). When $\gamma$ is a Gaussian process the second order statistics of the process $u(t)$ can also be found. Note, however, that the process $u(t)$ is in this case non-Gaussian since the integrals in (192) now contain terms $\exp(-\int \gamma(t')dt')$ which are not Gaussian.

- $\gamma(t)$ is given by a two-state Markov jump process: Particularly interesting dynamics can be obtained when only one of the states is positive but the mean damping $E[\gamma(t)]$ (see (182)) remains positive so that the dynamics is mean stable (some extra conditions have to be satisfied to assure mean stable dynamics but $E[\gamma(t)] > 0$ is a necessary condition [1]). Note that when the damping is negative the system (190) is in an unstable phase.
Figure 14: Two sample paths of the non-Gaussian complex scalar process $u(t)$ satisfying (190) and of the damping $\gamma$ represented by scalar the two-state Markov jump process with different ratios between the switching rates $\mu, \nu$. Left: a two state process with $\mu = 0.1$ (unstable to stable), $\nu = 0.05$ (stable to unstable). Right: $\mu = 0.8$, $\nu = 0.4$.

Since the damping in the model is piece-wise constant, the solution $u(t)$ can be computed on each interval where $\gamma$ is constant. In the stable regime

$$u(t) = e^{(-d^+ + i\omega)(t-t_0)}u_0 + \int_{t_0}^t f(s)e^{(-d^+ + i\omega)(t-s)}ds + \sigma \int_{t_0}^t e^{(-d^+ + i\omega)(t-s)}dW(s),$$  \hspace{1cm} (193)$$

whereas in the unstable regime

$$u(t) = e^{(-d^- + i\omega)(t-t_0)}u_0 + \int_{t_0}^t f(s)e^{(-d^- + i\omega)(t-s)}ds + \sigma \int_{t_0}^t e^{(-d^- + i\omega)(t-s)}dW(s).$$  \hspace{1cm} (194)$$

Figures 13-14 show a few typical examples of pathwise solutions of (190) when the damping $\gamma(t)$ is given by a two-state Markov jump process with different switching rates. Analytical formulas for the second order statistics of (190) can be found in [5].

4.2.2 Spatially extended turbulent system with transient instabilities

We finally introduce a relatively simple stochastic model PDE model [4] for a turbulent spatially extended system. Consider solutions of a real-valued stochastic partial differential equation (SPDE)

$$\frac{\partial u(x,t)}{\partial t} = \mathcal{P}\left(\frac{\partial}{\partial x}\right)u(x,t) - \Gamma\left(\frac{\partial}{\partial x}\right)u(x,t) + f(x,t) + \sigma(x) \dot{W}(t),$$  \hspace{1cm} (195)$$

where $\dot{W}$ is the white noise. Here, the initial data $u(x,t = t_0)$ is a Gaussian random field with nonzero covariance. Since we consider the problem in a one-dimensional periodic domain, the solution of (195) is given by the infinite Fourier series

$$u(x,t) = \sum_{k=-\infty}^{\infty} \hat{u}_k(t)e^{ikx}, \hspace{1cm} \hat{u}_k(t) = \hat{u}_k^*(t).$$  \hspace{1cm} (196)$$
The operators $\mathcal{P}$ and $\Gamma$ are defined at a given wavenumber by
\begin{align}
\mathcal{P} \left( \frac{\partial}{\partial x} \right) e^{ikx} &= \tilde{p}(ik)e^{ikx}, \\
\Gamma \left( \frac{\partial}{\partial x} \right) e^{ikx} &= \gamma(ik)e^{ikx}.
\end{align}

We assume that the spatially correlated noise can be represented as
\begin{equation}
\sigma(x) \dot{W}(t) = \sum_{k=-\infty}^{\infty} \sigma_k \dot{W}_k(t) e^{ikx},
\end{equation}
where $W_k(t)$ are independent complex Wiener processes for each $k \geq 0$, and their independent real and imaginary parts have the same variance $1/2$ and $\sigma_{-k} = \sigma_k$, $W_{-k} = W_k^*$. Substitution of (196)-(199) into (195) yields a system of uncoupled forced Langevin equations on each Fourier mode
\begin{equation}
d\hat{u}_k(t) = (\tilde{p}(ik) - \gamma(ik))\hat{u}_k(t)dt + \hat{f}_k(t)dt + \sigma_k dW_k(t), \quad \hat{u}_{-k}(t) = \hat{u}_k^*(t),
\end{equation}
where $\hat{f}_k(t)$ is the Fourier coefficient of the deterministic forcing $f(x,t)$. Note that (200) for each Fourier mode assumes the form (190) if we assume that $\tilde{p}(ik)$ is wave-like. i.e., when
\begin{equation}
\tilde{p}(ik) = i\omega_k,
\end{equation}
where $\omega_k$ is a real-valued, dispersion relation while $\gamma(ik)$ represents both explicit and turbulent dissipative processes. When $\gamma(ik) > 0$ is deterministic for all $k \neq 0$ Gaussian equilibrium distribution for (200) exists and, provided $f(x,t) = 0$, this statistical equilibrium distribution has mean zero and variance
\begin{equation}
E_k = \frac{\sigma_k^2}{2\gamma(ik)}, \quad 1 \geq k \geq \infty.
\end{equation}

As in the single mode case, one can induce regime switching in this spatially extended system by modeling the damping in a subset of the Fourier modes as a two-state Markov jump process. An example of such a situation is shown in figure 15; note the propagation of a coherent wave as a result of this transient instability.
Figure 15: Example of a transient instability in a spatially extended turbulent system induced by (temporarily) negative damping of some of the Fourier modes. The mode damping $\gamma_k$ is modelled as a two-state Markov jump process with a positive mean but only one of the states corresponding to (positive) damping.

References


