AM216 Stochastic Differential Equations

Lecture 06

List of topics in this lecture

- Energy spectrum density (ESD), power spectrum density (PSD)
- Stationary stochastic process, auto-correlation function (ACF)
- Wiener-Khinchin theorem: PSD is Fourier transform of ACF
- Definition of white noise: PSD is constant in frequency domain
- Calculating ACF and PSD of $Z(t) \equiv dW/dt$
- Constrained Wiener process, Bayes Theorem

Recap

Gambler's ruin problem:

Methodology of deriving BVPs for $u(x)$ and $T(x)$

Scaling and non-dimensionalization, advantage of working with $\dfrac{dW(t)}{\sqrt{dt}}$

Short story of white noise ...

Fourier transform: $F[y(t)] \equiv \int_{-\infty}^{+\infty} \exp(-i2\pi\xi t)\, y(t)\, dt$

Properties of Fourier transform (continued)

1) $F\left[\rho_{N(0,\sigma^2)}(t)\right] = F\left[\dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\dfrac{-t^2}{2\sigma^2}\right)\right] = \exp\left(-2\pi^2\sigma^2\xi^2\right)$

2) $F[\delta(x)] = 1$

3) $F[1] = \delta(\xi)$

---

4) Parseval's theorem

$$\int \left|y(t)\right|^2 dt = \int \left|\hat{y}(\xi)\right|^2 d\xi$$

Proof:

$$\int |y(t)|^2 dt = \int y(t)\overline{y(t)}dt = \int\left(\int \exp(i2\pi\xi t)\hat{y}(\xi)d\xi \int \exp(-i2\pi\eta t)\overline{\hat{y}(\eta)}d\eta\right)dt$$

$$= \int\left(\int\int \exp\left(-i2\pi(\eta-\xi)t\right)\hat{y}(\xi)\overline{\hat{y}(\eta)}d\xi d\eta\right)dt$$

Change the order of integration

$$= \int\int \hat{y}(\xi)\overline{\hat{y}(\eta)}\underbrace{\left(\int \exp\left(-i2\pi(\eta-\xi)t\right)dt\right)}_{F[1]=\delta(\eta-\xi)}d\eta d\xi$$

$$= \int\int \hat{y}(\xi)\overline{\hat{y}(\eta)}\delta(\eta-\xi)d\eta d\xi = \int \hat{y}(\xi)\overline{\hat{y}(\xi)}d\xi = \int |\hat{y}(\xi)|^2 d\xi$$

A rigorous proof: $\int |y(t)|^2 dt = \lim_{\sigma\to 0}\int y(t)\overline{y(t)}e^{-\sigma^2 t^2}dt = \cdots$

Recall the short story of white noise:

1) $Z(t) \equiv \dfrac{dW}{dt}$ is not a regular function.
2) $E\left(Z(t)Z(s)\right) = \delta(t-s)$
3) $\int \exp(-i2\pi\xi t)E\left(Z(t)Z(0)\right)dt = 1$
4) $Z(t)$ is a white noise (we will clarify what that means).

The long story of white noise

We follow the steps listed below.

- Energy $\propto \int_{-T}^{T}|y(t)|^2 dt$ --> Energy spectrum density (ESD)
- Power $\propto \dfrac{1}{T}\int_{-T}^{T}|y(t)|^2 dt$ --> Power spectrum density (PSD)
- Relation between PSD and auto-correlation function (ACF)
- Definition of white noise based on PSD
- Calculating ACF and PSD of $Z(t) \equiv dW/dt$

Energy spectrum density (ESD)

In many physics problems,

$$\text{Energy} \propto \int_{-\infty}^{+\infty} \left| y(t) \right|^2 dt$$

Examples:

$y(t)$ = electric current

$$\text{Energy} = \int_{-\infty}^{+\infty} R \cdot y(t)^2 \, dt\,, \qquad R = \text{electrical resistance}$$

Here "energy" refers to the dissipated energy.

$y(t)$ = velocity

$$\text{Energy} = \int_{-\infty}^{+\infty} b \cdot y(t)^2 \, dt\,, \qquad b = \text{viscous drag coefficient}$$

For mathematical convenience, we define

$$\text{Energy} \equiv \underbrace{\int_{-\infty}^{+\infty} \left| y(t) \right|^2 dt = \int_{-\infty}^{+\infty} \left| \hat{y}(\xi) \right|^2 d\xi}_{\text{Parseval's theorem}} = \int_{-\infty}^{+\infty} \left| \int_{-\infty}^{+\infty} \exp(-i2\pi\xi t) y(t) dt \right|^2 d\xi$$

We like to know how the energy is distributed in the frequency domain.

Definition of energy spectrum density (ESD)

$$\text{ESD} \equiv \left| \hat{y}(\xi) \right|^2 = \left| \int_{-\infty}^{+\infty} \exp(-i2\pi\xi t) y(t) dt \right|^2$$

Caution: $\left| \hat{y}(\xi) \right|^2$ is an unnormalized density.

$$\int_{-\infty}^{+\infty} \left| \hat{y}(\xi) \right|^2 d\xi = \int_{-\infty}^{+\infty} \left| y(t) \right|^2 dt = \text{Energy} \neq 1$$

Other examples of unnormalized density:

Population density: X number of persons per square mile

Pollution density: X amount of chemicals per unit volume of air or water

Car density: X number of cars per mile of highway

Caution: (slightly different definitions of ESD)

In electrical engineering (EE), energy spectrum density is defined as

$$\text{ESD} \equiv \Phi(\omega) = \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp(-i\omega t) y(t) dt \right|^2 \qquad \omega\text{: angular frequency}$$

$\Phi(\omega)$ and $\left| \hat{y}(\xi) \right|^2$ are related by:

$$\Phi(\omega)=\frac{1}{2\pi}\left|\hat{y}(\xi)\right|^2, \qquad \xi=\frac{\omega}{2\pi}: \text{ ordinary frequency}$$

Power spectrum density (PSD)

Energy spectrum density is meaningful only when $\int_{-\infty}^{+\infty}\left|y(t)\right|^2 dt=\text{finite}$.

Example: $y(t)$ = electric current = y_0 = constant in time

$$\text{Energy}=\int_{-\infty}^{+\infty}R\cdot y_0^2\,dt=\infty$$

When the total energy is unbounded, we look at the energy per time.

$$\int_{-T}^{T}\left|y(t)\right|^2 dt=\int_{-\infty}^{\infty}\left|\int_{-T}^{T}\exp(-i2\pi\xi t)y(t)dt\right|^2 d\xi \qquad \text{(Parseval's theorem)}$$

$$\frac{1}{2T}\int_{-T}^{T}\left|y(t)\right|^2 dt=\int_{-\infty}^{\infty}\frac{1}{2T}\left|\int_{-T}^{T}\exp(-i2\pi\xi t)y(t)dt\right|^2 d\xi$$

Definition of power spectrum density (PSD)

$$\text{PSD}\equiv\lim_{T\to\infty}\frac{1}{2T}\left|\int_{-T}^{T}\exp(-i2\pi\xi t)y(t)dt\right|^2$$

Expression of power spectrum density (PSD)

We write PSD into a more workable expression.

$$\text{PSD}=\lim_{T\to\infty}\frac{\int_{-T}^{T}\exp(-i2\pi\xi t)y(t)dt\int_{-T}^{T}\exp(i2\pi\xi s)\overline{y(s)}ds}{2T}$$

$$=\lim_{T\to\infty}\frac{\int_{-T}^{T}\int_{-T}^{T}\exp(-i2\pi\xi(t-s))y(t)\overline{y(s)}dt\,ds}{2T}$$

change of variable $\tau = t - s$

$$=\lim_{T\to\infty}\frac{\int_{-T}^{T}\int_{-T-s}^{T-s}\exp(-i2\pi\xi\tau)y(\tau+s)\overline{y(s)}d\tau\,ds}{2T}$$

Draw the integration region in s-τ plane.

For each s, the range for τ is $[-T-s, T-s]$.

For each τ, the range for s is $[a(\tau), b(\tau)]$ where

$$a(\tau)=\begin{cases}-T-\tau, & \tau\in[-2T,\,0]\\ -T, & \tau\in[0,\,2T]\end{cases}, \qquad b(\tau)=\begin{cases}T, & \tau\in[-2T,\,0]\\ T-\tau, & \tau\in[0,\,2T]\end{cases}$$

Change the order of integration

$$\text{PSD} = \lim_{T\to\infty} \frac{\int_{-2T}^{2T} \exp(-i2\pi\xi\tau) \int_{a(\tau)}^{b(\tau)} y(\tau+s)\overline{y(s)}\, ds\, d\tau}{2T} \tag{PSD01}$$

So far, we worked with deterministic process $y(t)$.

Next we introduce stochastic process and stationary stochastic process.

Definition of stochastic process

A stochastic process is a function of time that varies with ω.

$$\underbrace{y(t)}_{\text{Short notation}} = \underbrace{y(t,\omega)}_{\text{Full notation}} \qquad \omega = \text{random outcome of an experiment}$$

Definition of stationary stochastic process

Let $y(t)$ be a stochastic process. We say $y(t)$ is stationary if for any set of time instances $(t_1, t_2, \ldots, t_k)$, the joint distribution of $(y(t+t_1), y(t+t_2), \ldots, y(t+t_k))$ is independent of t.

Examples:

- $W(t)$ is a stochastic processes. It is not stationary:
  $\text{var}(W(t)) = t$ varies with t.
- $Z(t) = \dfrac{dW(t)}{dt} \sim \dfrac{1}{\sqrt{dt}} N(0, dt)$ is a well defined stochastic process for finite dt.

  It is stationary: the joint distribution is invariant under a shift.

Properties of stationary stochastic process

For a stationary stochastic process, we have

- $E(y(t)) = E(y(0))$
- $\text{var}(y(t)) = \text{var}(y(0))$
- $E\left(y(s+\tau)\overline{y(s)}\right) = E\left(y(\tau)\overline{y(0)}\right)$

Definition of auto-correlation function (ACF)

For a stationary stochastic process $y(t)$, the auto-correlation function (ACF) is

$$R(\tau) \equiv E\left(y(s+\tau)\overline{y(s)}\right) = E\left(y(\tau)\overline{y(0)}\right)$$

Note: $R(\tau)$ is independent of s (for a stationary process).

Caution: be careful with the term "auto-correlation"

Auto-correlation coefficient is defined as

$$\rho(\tau) \equiv \frac{E\left(\left[y(\tau)-E(y(0))\right]\left[\overline{y(0)-E(y(0))}\right]\right)}{\mathrm{var}(y(0))}$$

Auto-correlation function (ACF) is defined as

$$R(\tau) \equiv E\left(y(\tau)\overline{y(0)}\right)$$

Relation between PSD and ACF

For a stationary stochastic process, the power spectrum density (PSD) is

$$\underbrace{s(\xi)}_{\substack{\text{New notation} \\ \text{for PSD}}} \equiv \mathrm{PSD} \equiv \lim_{T\to\infty} \frac{E\left(\left|\int_{-T}^{T} \exp(-i2\pi\xi t) y(t)\, dt\right|^2\right)}{2T}$$

We use (PSD01), obtained above for a deterministic process, to rewrite $s(\xi)$

$$\mathrm{s}(\xi) = \lim_{T\to\infty} \frac{E\left(\int_{-2T}^{2T} \exp(-i2\pi\xi\tau) \int_{a(\tau)}^{b(\tau)} y(\tau+s)\overline{y(s)}\, ds\, d\tau\right)}{2T}$$

Change the order of integration and expectation

$$= \lim_{T\to\infty} \frac{\int_{-2T}^{2T} \exp(-i2\pi\xi\tau) \int_{a(\tau)}^{b(\tau)} E\left(y(\tau+s)\overline{y(s)}\right) ds\, d\tau}{2T}$$

$$= \lim_{T\to\infty} \frac{\int_{-2T}^{2T} \exp(-i2\pi\xi\tau) \int_{a(\tau)}^{b(\tau)} R(\tau)\, ds\, d\tau}{2T} \qquad R(\tau) \text{ is independent of } s.$$

$$= \lim_{T\to\infty} \frac{\int_{-2T}^{2T} \exp(-i2\pi\xi\tau) R(\tau)(b(\tau)-a(\tau))\, d\tau}{2T}$$

The term $(b(\tau) - a(\tau))$ has the expression:

$$b(\tau)-a(\tau) = \begin{cases} 2T+\tau, & \tau \in [-2T,\ 0] \\ 2T-\tau, & \tau \in [0,\ 2T] \end{cases} = 2T - |\tau|$$

Substituting it into the expression of $s(\xi)$ yields

$$s(\xi) = \lim_{T\to\infty} \int_{-2T}^{2T} \exp(-i2\pi\xi\tau) R(\tau) \left(1 - \frac{|\tau|}{2T}\right) d\tau$$

Taking the limit as $T \to \infty$, we arrive at

$$s(\xi)=\int_{-\infty}^{+\infty}\exp(-i2\pi\xi\tau)R(\tau)d\tau$$

We just derived the Wiener-Khinchin theorem.

Wiener-Khinchin theorem (relation between PSD and ACF)

For a stationary stochastic process $y(t)$, the power spectrum density, $s(\xi)$, and the auto-correlation function, $R(t)$, are related by

$$s(\xi)=\int_{-\infty}^{+\infty}\exp(-i2\pi\xi t)R(t)dt$$

In other words, the PSD is the Fourier transform of ACF.

Definition of white noise

Let $y(t)$ be a stationary stochastic process. We say $y(t)$ is a white noise if

$s(\xi) = const$

In other words, the power of a white noise is uniformly distributed in the frequency domain. The Wiener-Khinchin theorem tells us that

$$s(\xi)=const \qquad \Longleftrightarrow \qquad R(t)\equiv E\left(y(t)\overline{y(0)}\right)\propto\delta(t)$$

Working out items in the short story

We re-write the short story in terms of the auto-correlation function $R(\tau)$ and power spectrum density $s(\xi)$.

1) $Z(t)\equiv\dfrac{dW}{dt}$ is not a regular function.

2) $R(\tau)=E\left(Z(s+\tau)Z(s)\right)=\delta(\tau)$

3) $s(\xi)=\int\exp(-i2\pi\xi t)R(t)dt=1$

4) $Z(t)$ is a white noise.

- To show $Z(t)$ is a white noise (item 4), we only need $s(\xi)$ = const (item 3).
- To show $s(\xi) = 1$(item 3), we only need $R(t) = \delta(t)$ (item 2)

Thus, the remaining task is to show item 2, which we do now.

Derivation of $R(t) = \delta(t)$ for $Z(t) \equiv dW/dt$

Here we present a "formal" derivation. A rigorous derivation is in Appendix A.

We first calculate $E(W(t)W(s))$ for $t \ge s$.

$$E\left(W(t)W(s)\right)=E\left((W(t)-W(s)+W(s))W(s)\right)$$

$$= E\Big((W(t)-W(s))W(s)\Big)+E\Big(W(s)^2\Big)=0+s=s$$

Since $E(W(t)W(s)) = E(W(s)W(t))$, we obtain

$$E\Big(W(t)W(s)\Big)=\min(t,s)$$

Next, in the calculation of $E(Z(t)Z(s))$, we "formally" exchange the order of differentiation and expectation.

$$E\Big(Z(t)Z(s)\Big)=E\left(\frac{\partial}{\partial s}\frac{\partial}{\partial t}\Big(W(t)W(s)\Big)\right)$$

$$=\frac{\partial}{\partial s}\frac{\partial}{\partial t}E\Big(W(t)W(s)\Big)=\frac{\partial}{\partial s}\frac{\partial}{\partial t}\min(t,s)$$

As a function of t, we have

$$\min(t,s)=\begin{cases} t, & t<s \\ s, & t>s \end{cases}$$

Differentiating with respect to t, and then writing it as a function of s, we get

$$\frac{\partial}{\partial t}\min(t,s)=\begin{cases} 1, & t<s \\ 0, & t>s \end{cases} \qquad \text{(as a function of } t\text{)}$$

$$=\begin{cases} 0, & s<t \\ 1, & s>t \end{cases} \qquad \text{(as a function of } s\text{)}$$

Differentiating with respect to s, we arrive at

$$\frac{\partial}{\partial s}\frac{\partial}{\partial t}\min(t,s)=\delta(s-t)$$

Therefore, we conclude

$$E\Big(Z(t)Z(s)\Big)=\frac{\partial}{\partial s}\frac{\partial}{\partial t}\min(t,s)=\delta(s-t)$$

$$\Longrightarrow \qquad R(\tau)=E\Big(Z(s+\tau)Z(s)\Big)=\delta(\tau)$$

<u>A class of colored noise</u>:

In the subsequent discussion of Ornstein-Uhlenbeck process (OU), we will see that its auto-correlation has the form:

$$R(t) = E\left(y(t)\overline{y(0)}\right) \propto \exp(-\beta|t|)$$

The corresponding power spectrum density is

$$s(\xi) = \int \exp(-i2\pi\xi t)R(t)dt \propto \int \exp(-i2\pi\xi t)\exp(-\beta|t|)dt = \frac{2\beta}{\beta^2 + 4\pi^2\xi^2}$$

End of discussion of white noise

Constrained Wiener process

For an unconstrained Wiener process, we have

$W(0) = 0$ and $W(t_1) \sim N(0, t_1)$

Question: What happens if it is constrained by $W(t_1+t_2) = y$?

We like to know the conditional distribution $(W(t_1) \mid W(t_1+t_2) = y)$.

To answer this question, we need to introduce Bayes theorem.

Bayes Theorem

Consider two events A and B. We write Pr(A and B) in two ways.

Pr(A and B) = Pr(A | B) Pr(B)

Pr(A and B) = Pr(B | A) Pr(A)

Equating the two, we get

Pr(A | B) Pr(B) = Pr(B | A) Pr(A)

Express Pr(A | B) in terms of Pr(B | A), we arrive at

Bayes Theorem for events:

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B)}$$

To derive Bayes theorem for densities, we consider

A = "$x < X \leq x+\Delta x$"

B = "$y < Y \leq y+\Delta y$"

We write probabilities in terms of densities

$$\Pr(A|B) \approx \rho(X = x|Y = y)\Delta x$$

$$\Pr(B|A) \approx \rho(Y = y|X = x)\Delta y$$

$\Pr(A) \approx \rho(X = x)\Delta x$

$\Pr(B) \approx \rho(Y = y)\Delta y$

Substituting these terms into Bayes theorem, we obtain…

Bayes theorem for densities

$$\rho(X = x \mid Y = y) = \frac{\rho(Y = y \mid X = x)\cdot\rho(X = x)}{\rho(Y = y)}$$

A useful trick:

In density $\rho(X=x|Y=y)$, x is the independent variable and y is a parameter. On the RHS of Bayes theorem, $\rho(Y=y)$ has no dependence on x and serves as a normalizing factor.

Thus, we don't need to explicitly keep track of $\rho(Y=y)$. We can write Bayes theorem conveniently as

$$\boxed{\rho(X = x \mid Y = y) \propto \rho(Y = y \mid X = x)\cdot\rho(X = x)}$$

where the RHS needs a proper normalizing factor to make it integrate to 1.

This trick is especially convenient for normal distributions:

$$X \sim N(\mu, \sigma^2) \qquad \Longleftrightarrow \qquad \rho(X = x) \propto \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right),$$

Conditional density $\rho(W(t_1) = x \mid W(t_1+t_2) = y)$

To use the Bayes theorem, we first find $\rho(W(t_1) = x)$ and $\rho(W(t_1+t_2) = y \mid W(t_1) = x)$.

$$W(t_1) \sim N(0, t_1)$$

$$\Longrightarrow \quad \rho\big(W(t_1) = x\big) = \rho_{N(0,t_1)}(x) \propto \exp\left(\frac{-x^2}{2t_1}\right)$$

$$W(t_1+t_2) = W(t_1) + \underbrace{\big(W(t_1+t_2) - W(t_1)\big)}_{\sim N(0,t_2)}$$

$$\Longrightarrow \quad \big(W(t_1+t_2) \mid W(t_1) = x\big) \sim N(x, t_2)$$

$$\Longrightarrow \quad \rho\big(W(t_1+t_2) = y \mid W(t_1) = x\big) = \rho_{N(x,t_2)}(y) \propto \exp\left(\frac{-(y-x)^2}{2t_2}\right)$$

The Bayes theorem gives us

$$\rho\Big(W(t_1)=x \,\Big|\, W(t_1+t_2)=y\Big) \propto \rho\Big(W(t_1+t_2)=y \,\Big|\, W(t_1)=x\Big)\cdot\rho\Big(W(t_1)=x\Big)$$

(We don't need to keep track of factors that are independent of x!)

$$\propto \exp\left(\frac{-(y-x)^2}{2t_2}\right)\exp\left(\frac{-x^2}{2t_1}\right) \propto \exp\left(-\left(\left(\frac{1}{2t_1}+\frac{1}{2t_2}\right)x^2 - 2\frac{y}{2t_2}x\right)\right)$$

(Completing the square)

$$\propto \exp\left(\frac{-\left(x-\frac{t_1 y}{t_1+t_2}\right)^2}{2\frac{t_1 t_2}{t_1+t_2}}\right) \sim N\left(\frac{t_1 y}{t_1+t_2}, \frac{t_1 t_2}{t_1+t_2}\right)$$

We conclude

$$\boxed{\rho\Big(W(t_1)=x \,\Big|\, W(t_1+t_2)=y\Big) \sim N\left(\frac{t_1 y}{t_1+t_2}, \frac{t_1 t_2}{t_1+t_2}\right)}$$

<u>For the general case</u>, we have

$$\boxed{\rho\Big(W(a+t_1)=x \,\Big|\, W(a)=y_a \text{ and } W(a+t_1+t_2)=y_b\Big) \sim N\left(\frac{t_1 y_b + t_2 y_a}{t_1+t_2}, \frac{t_1 t_2}{t_1+t_2}\right)}$$

<u>A special case:</u> $t_1 = t_2 = h/2$

$$\rho\left(W(a+\frac{h}{2})=x \,\Big|\, W(a)=y_a \text{ and } W(a+h)=y_b\right) \sim N\left(\frac{y_a+y_b}{2}, \frac{h}{4}\right)$$

This is very useful in refining a discrete sample path of $W(t)$.

Appendix A: A rigorous derivation of $R(t)$ and $s(\xi)$ for $Z(t) \equiv dW/dt$

First, we work with finite dt. Let $\Delta t \equiv dt$. We have

$$Z(t) = \frac{W(t+\Delta t)-W(t)}{\Delta t} \qquad \text{a well defined stationary stochastic process}$$

$$E\big(Z(t)Z(s)\big) = E\left(\frac{W(t+\Delta t)-W(t)}{\Delta t}\cdot\frac{W(s+\Delta t)-W(s)}{\Delta t}\right)$$

$$= \begin{cases} 0, & |t-s| > \Delta t \\ \dfrac{\Delta t - |t-s|}{(\Delta t)^2}, & |t-s| \le \Delta t \end{cases} \qquad \text{(derivation not included)}$$

$$R(\tau) = E\big(Z(s+\tau)Z(s)\big) = \begin{cases} 0, & |\tau| > \Delta t \\ \dfrac{\Delta t - |\tau|}{(\Delta t)^2}, & |\tau| \le \Delta t \end{cases}$$

Taking the Fourier transform of $R(t)$, we obtain

$$s(\xi) = \int \exp(-i2\pi\xi t)R(t)\,dt = \int_{-\Delta t}^{\Delta t} \exp(-i2\pi\xi t)\frac{\Delta t - |t|}{(\Delta t)^2}\,dt$$

$$= 2\frac{\cosh(i2\pi\xi\Delta t)-1}{(i2\pi\xi\Delta t)^2} \qquad \text{(derivation not included)}$$

Finally, we take the limit as $\Delta t \to 0$. At any fixed ξ, as $\Delta t \to 0$, we have

$$\lim_{\Delta t\to 0} s(\xi) = \lim_{\Delta t \to 0} 2\frac{\cosh(i2\pi\xi\Delta t)-1}{(i2\pi\xi\Delta t)^2} = 1$$

Observation:

- Mathematically, working with finite dt until taking the limit at the end is a rigorous approach in which every step is properly justified.
- "Formal" derivations are not rigorous but are much simpler.