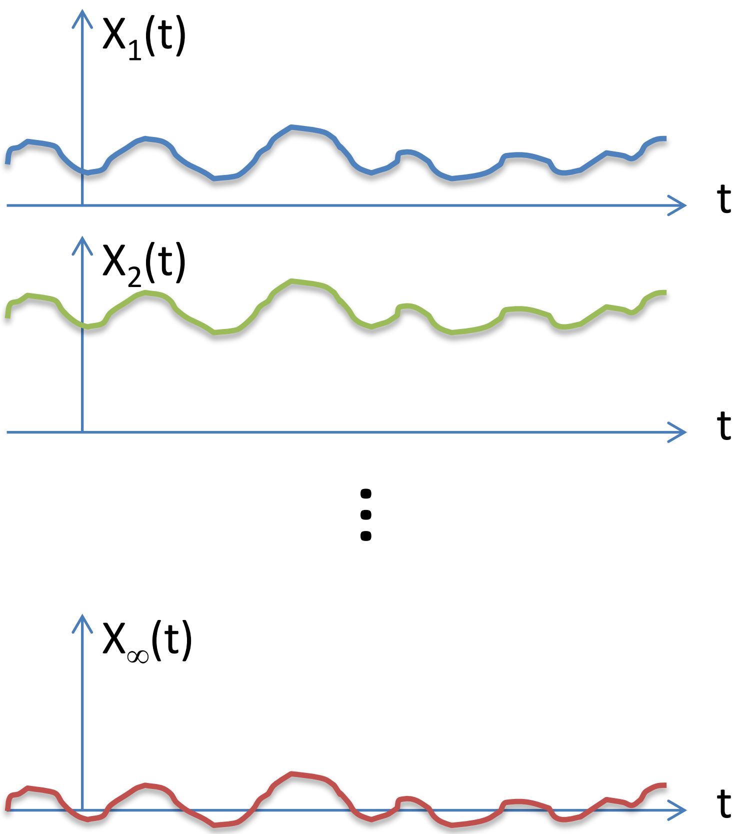
Autokorrelation og krydskorrelation

Læsning:

Cooper/McGillem kap. 6

Hvad er en stokastisk proces?

- Et *ensemble* af tidslige funktioner og den tilhørende tætheds- eller fordelingsfunktion.



Hver af de mulige tidsfunktioner i ensemblet, repræsenterer et specifikt valg af *stokastiske* procesparametre.

I eksemplet til venstre har jeg forsøgt at skitsere en stokastisk proces, hvor den eneste procesparameter er offset'et på en vertikale akse.

De stokastiske procesparametre har en tilhørende tæthedsfunktion. Fx kunne offset'et i dette eksempel være uniformt fordelt.

Stationær i bred forstand

- Hvis der for alle valg af t_1 og t_2 gælder

$$E\big(X(t_1)\big) = E\big(X(t_2)\big)$$

- og der for alle valg af t_1 og t_2

$$E\big(X(t_1) \cdot X(t_2)\big) = funktion(t_1 - t_2)$$

- Middelværdien skal være konstant, og korrelationen mellem $X(t_1)$ og $X(t_2)$ skal være en funktion af tidsforskellen $t_1 - t_2$.

Ergodisk proces

- Hvis ensemble midling = tidslig midling

$$\bar{X} = \int_{-\infty}^{\infty} x \cdot f(x) dx = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, dt$$

- Helt generelt skal der gælde, at

$$\overline{X^n} = \int_{-\infty}^{\infty} X^n \cdot f(x) dx = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x^n(t)\, dt$$

Samplede/diskrete data

- Fra sidste forelæsning:

$$\hat{\overline{\overline{X}}} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

- Hvad er middelværdien af estimatet ($\hat{\overline{\overline{X}}}$) af middelværdien?
- Unbiased, dvs.

$$E[\hat{\overline{\overline{X}}}] = \overline{X}$$

Intuition om korrelation

- Autokorrelationen $E[X(t) \cdot X(t+\tau)]$
  - siger noget om, hvor meget signalet $X(t)$ ligner sig selv til tiden $X(t+\tau)$.
  - må afhænge af, hvor hurtigt signalet $X(t)$ ændrer sig over tid.
  - må være stor, hvis τ er lille.

- Krydskorrelationen $E[X(t) \cdot Y(t+\tau)]$
  - kan bruges til at lede efter steder (tidspunkter τ), hvor signalet $X(t)$ ligner signalet $\mathrm{Y}(t)$.

Autokorrelation

- Generelt

$$R_X(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = E[X_1 \cdot X_2]$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 \cdot x_2 \cdot f(x_1, x_2)\, dx_1 dx_2$$

- For en stationær proces

$$R_X(t_1, t_2) = R_X(t_1 + T, t_2 + T) = E[X(t_1 + T) \cdot X(t_2 + T)]$$

eller bare $R_X(\tau) = E[X(t) \cdot X(t + \tau)]$

Tidslig autokorrelation

- Generelt

$$\mathcal{R}_X(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot x(t+\tau) d\tau$$

- Hvis processen er ergodisk

$$R_X(\tau) = \mathcal{R}_X(\tau)$$

Effekt af τ

- $\tau = 0$:

$$R_X(0) = E[X(t) \cdot X(t+0)] = E\left[\left(X(t)\right)^2\right]$$ = ”mean square”

- $\tau \neq 0$:

$R_X(\tau) = E[X(t) \cdot X(t+\tau)]$ er et udtryk for similariteten mellem $X(t)$ og $X(t+\tau)$

Eksempel

- Hvor meget indgår $X(t+\tau)$ i $X(t)$

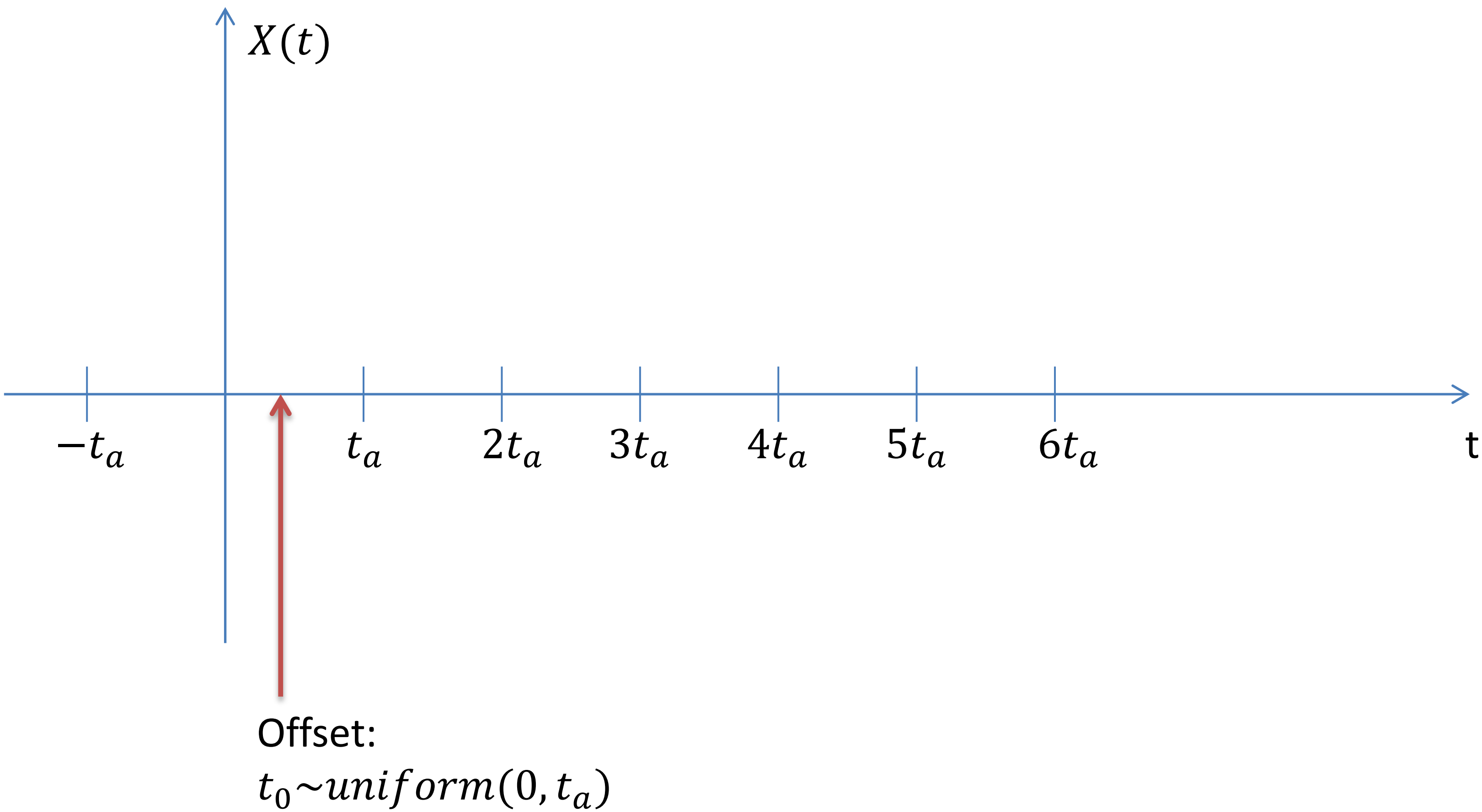
$$Y(t) = X(t) - \rho X(t+\tau)$$

hvor $\mathrm{E}[X(t)] = 0$.

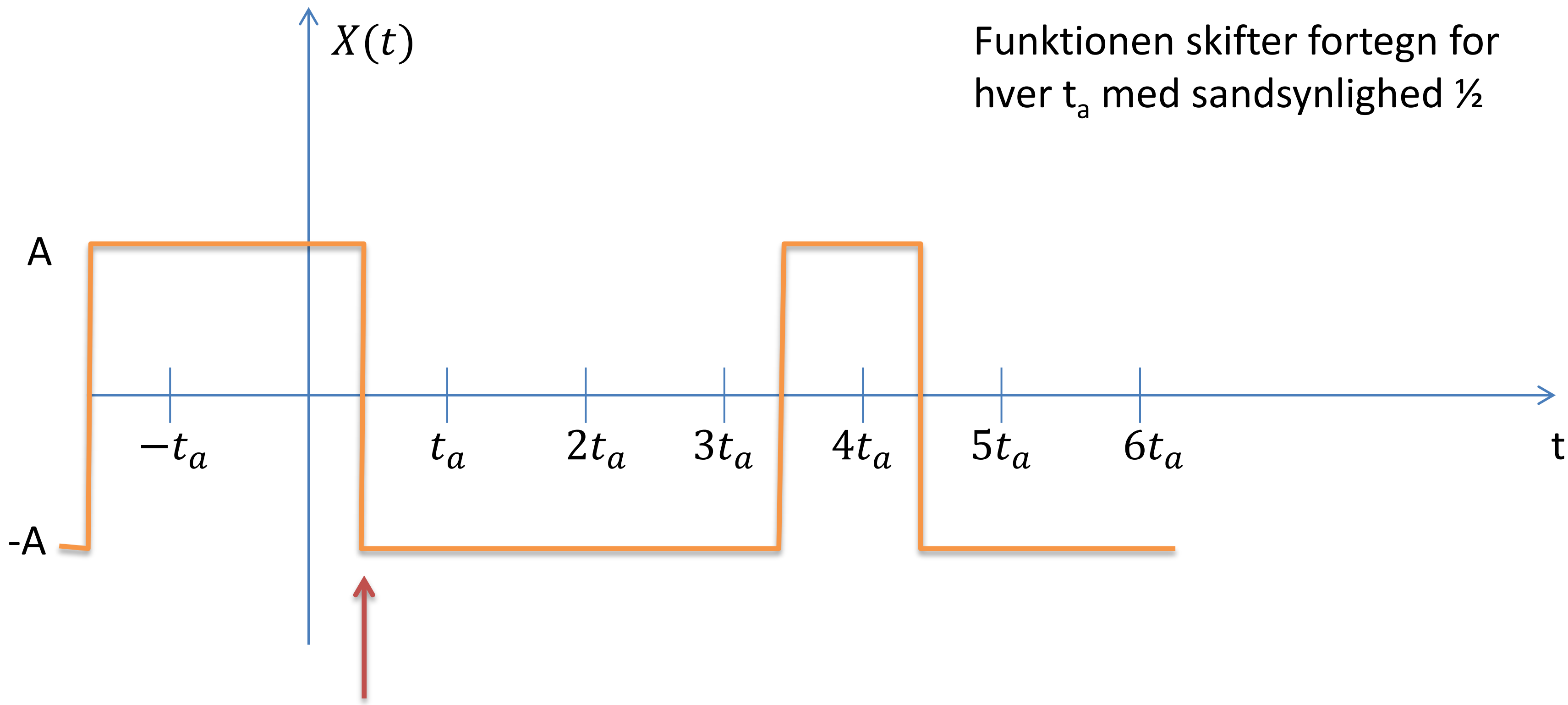
- Parameteren ρ måler, hvor meget $X(t+\tau)$ i $X(t)$.
- Den optimale værdi for ρ er den, hvor $\mathrm{E}[\left(Y(t)\right)^2]$ er mindst.
- Den optimale ρ er præcis korrelationskoefficienten, og er

$$\rho = \frac{R_X(\tau)}{\sigma_X^2}$$

Eksempel: Figur 6-1



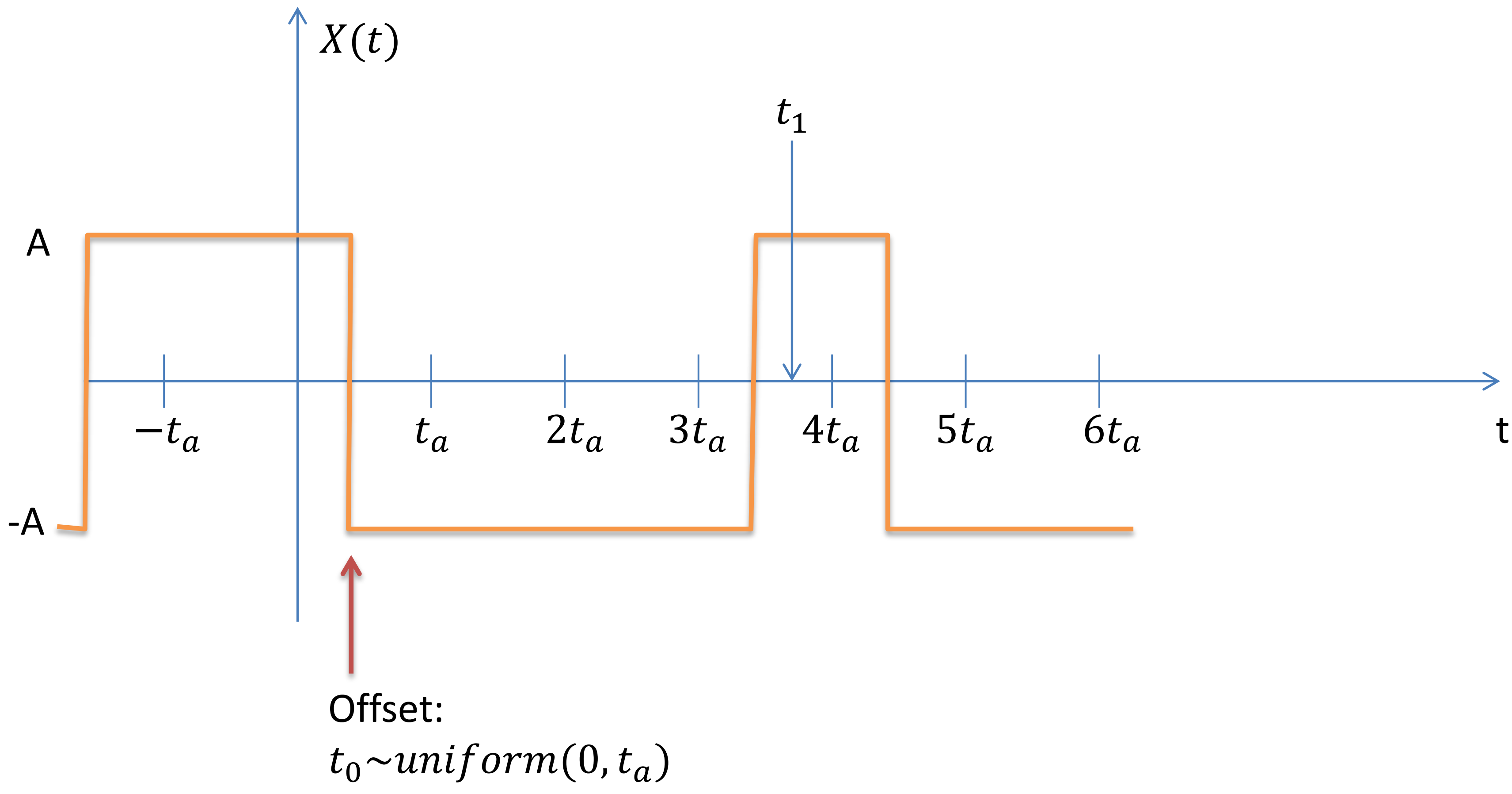
Eksempel: Figur 6-1



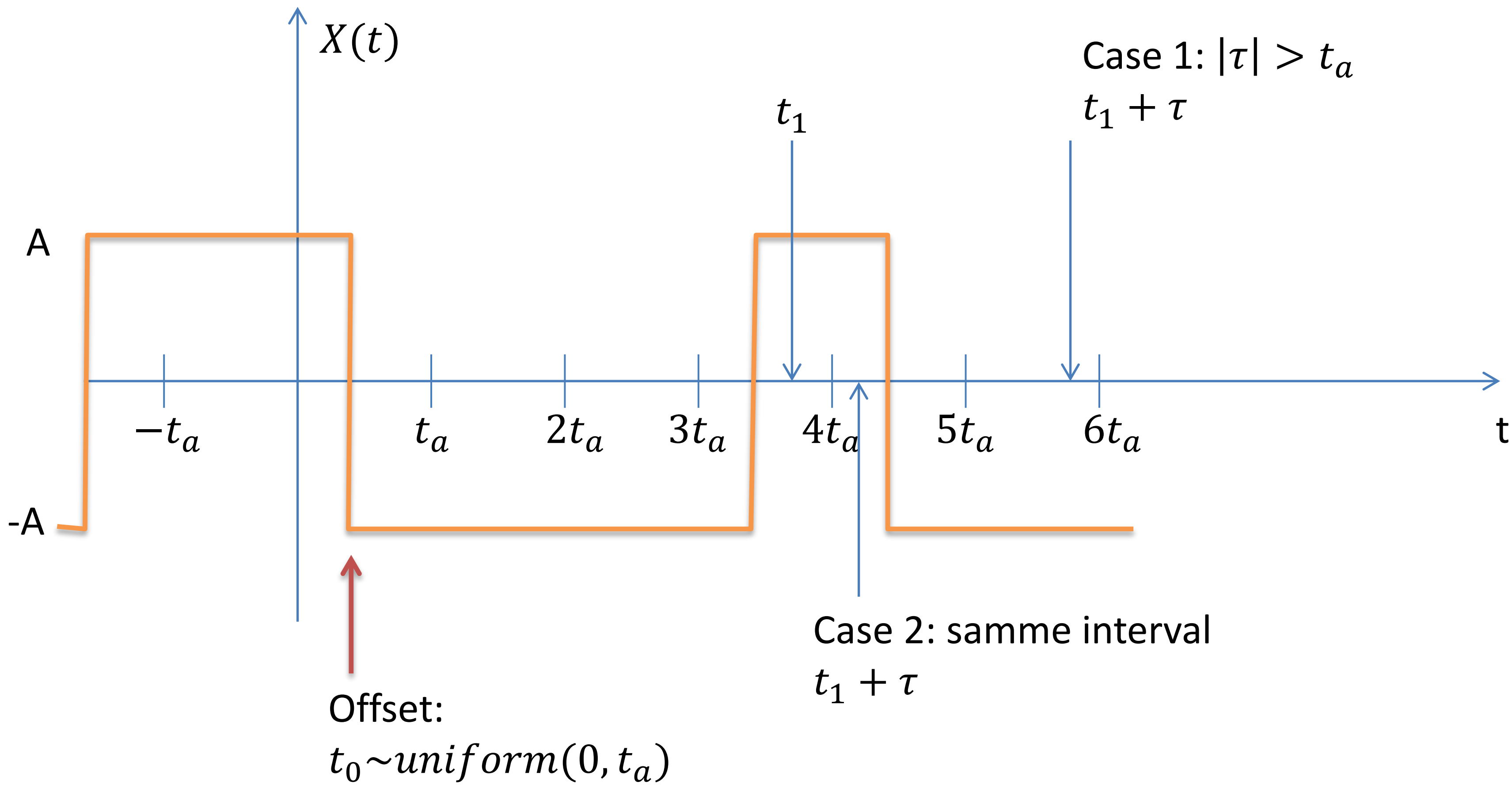
Funktionen skifter fortegn for hver t_a med sandsynlighed ½

Offset:
$t_0 \sim uniform(0, t_a)$

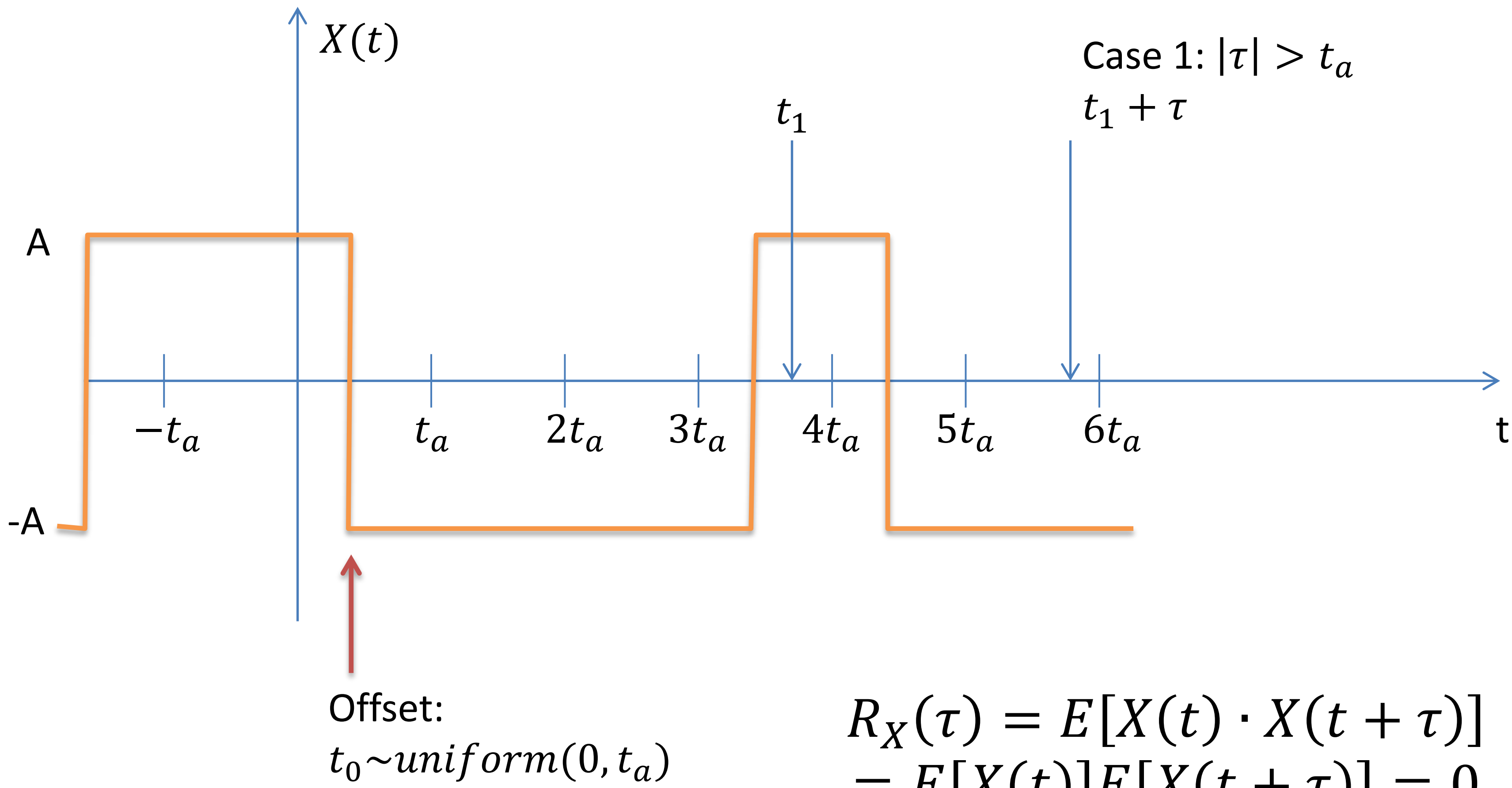
Eksempel: Figur 6-1



Eksempel: Figur 6-1

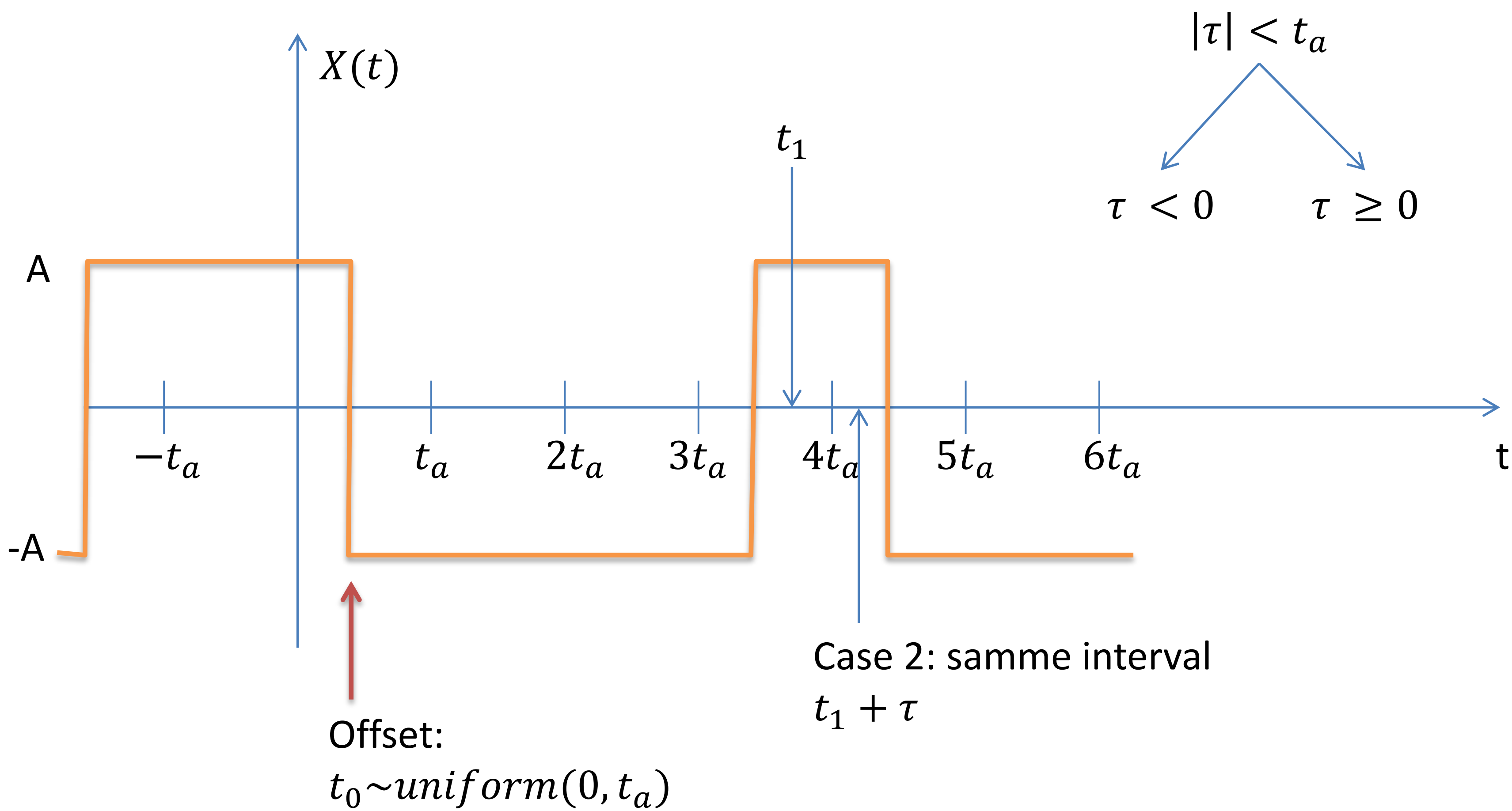


Eksempel: Figur 6-1

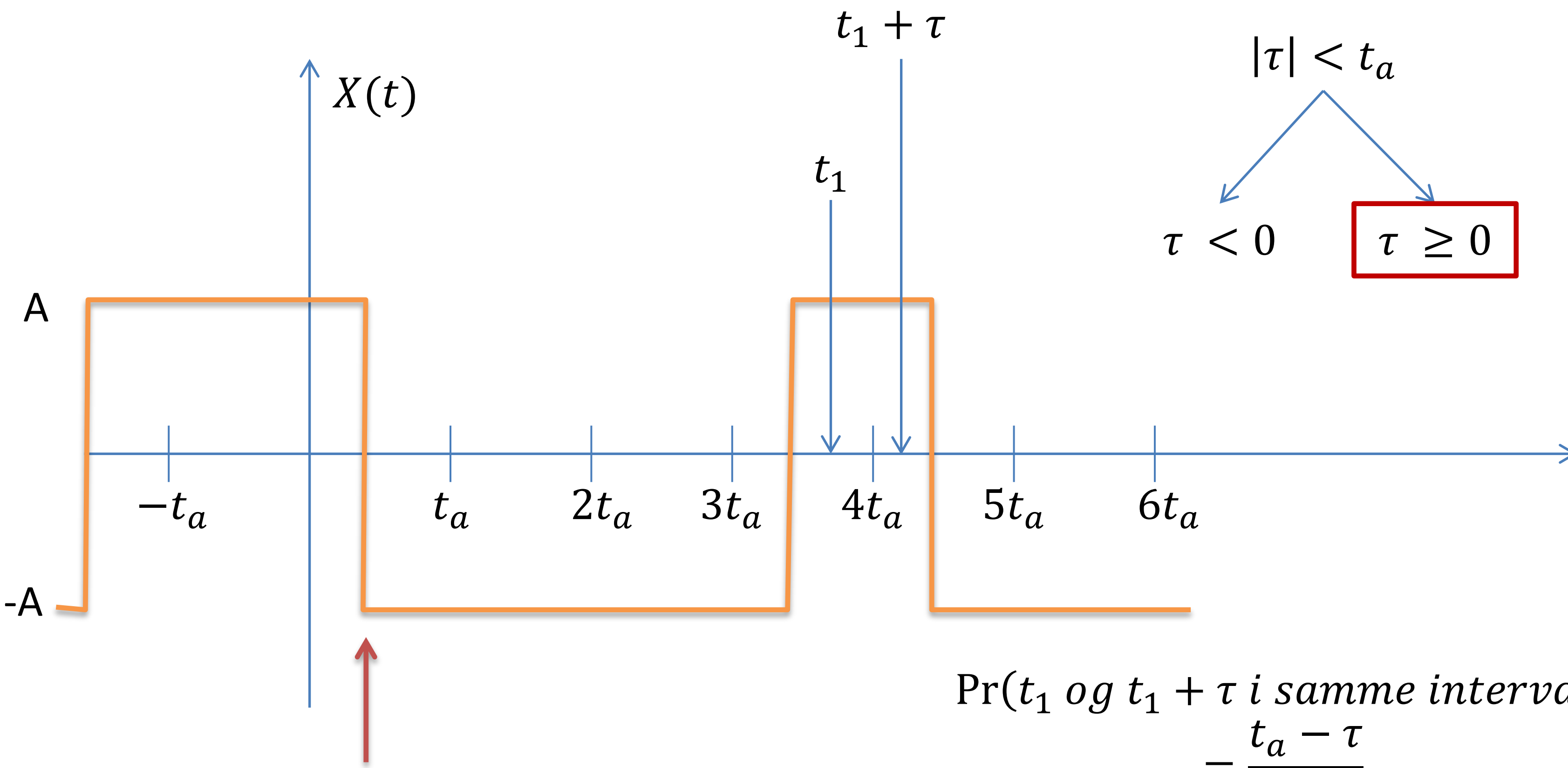


$$R_X(\tau) = E[X(t) \cdot X(t+\tau)]$$
$$= E[X(t)]E[X(t+\tau)] = 0$$

Eksempel: Figur 6-1



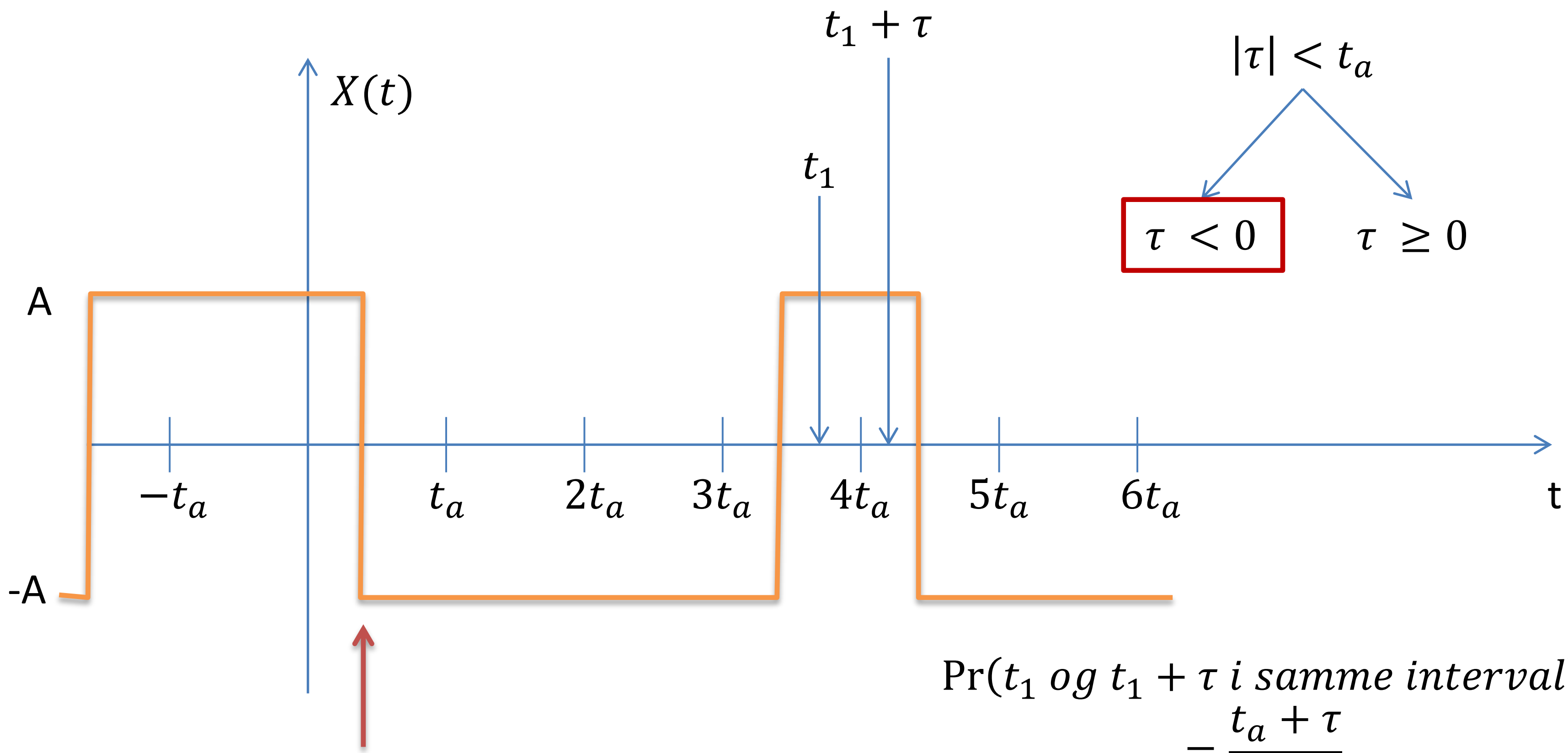
Eksempel: Figur 6-1



Offset:
$t_0 \sim uniform(0, t_a)$

$$\Pr(t_1 \ og \ t_1 + \tau \ i \ samme \ interval) = \frac{t_a - \tau}{t_a}$$

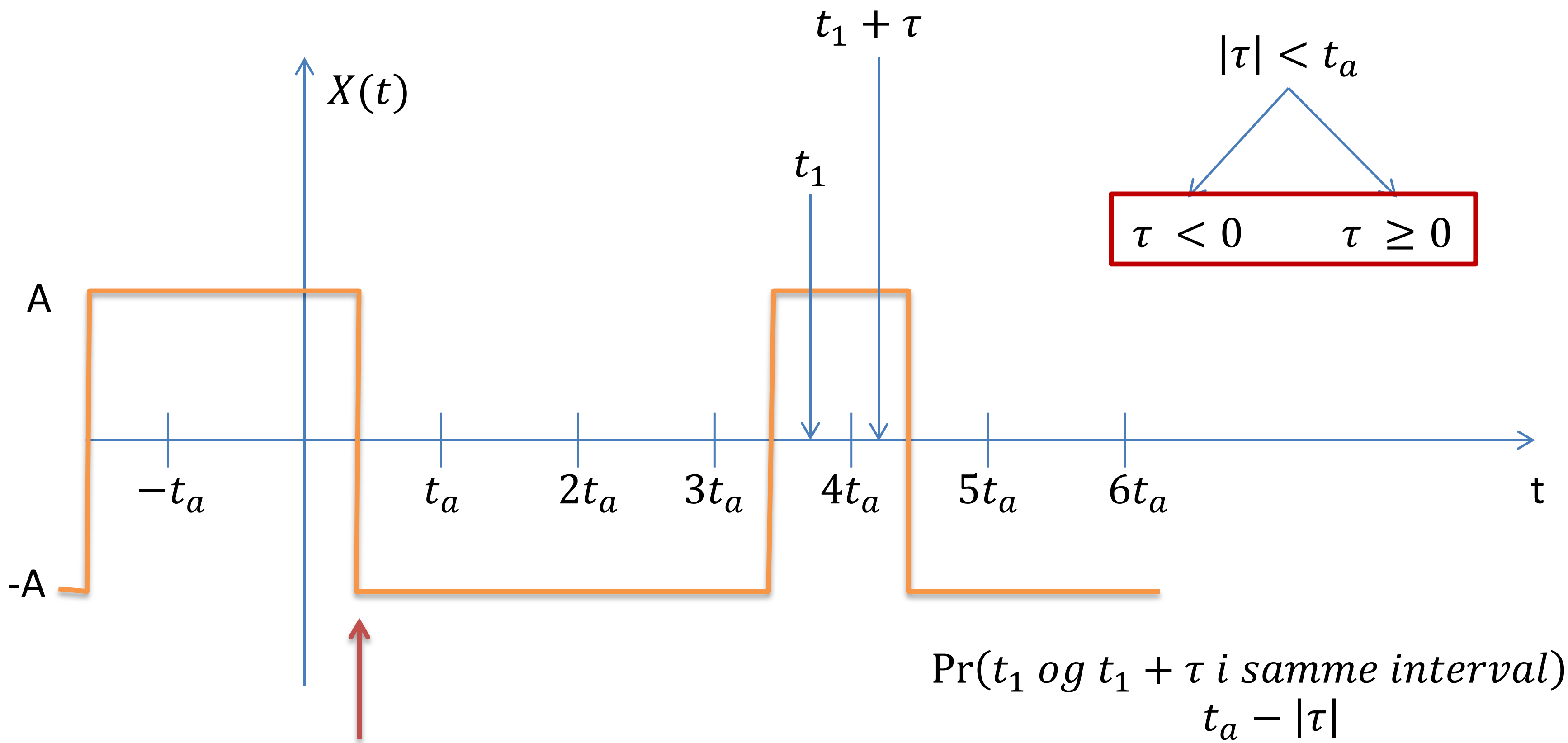
Eksempel: Figur 6-1



$$\Pr(t_1 \; og \; t_1 + \tau \; i \; samme \; interval) = \frac{t_a + \tau}{t_a}$$

Offset:
$t_0 \sim uniform(0, t_a)$

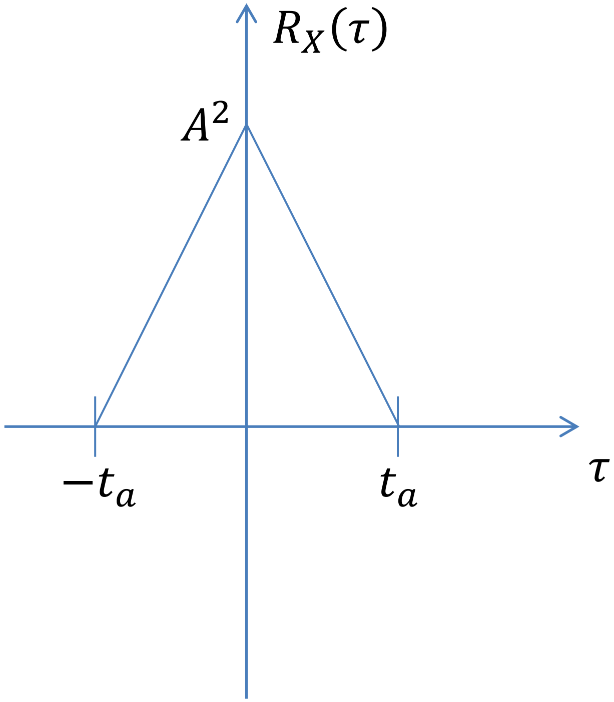
Eksempel: Figur 6-1



Offset:
$t_0 \sim uniform(0, t_a)$

$$\Pr(t_1 \; og \; t_1 + \tau \; i \; samme \; interval) = \frac{t_a - |\tau|}{t_a}$$

Eksempel: Figur 6-1

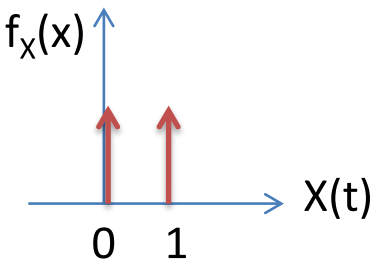


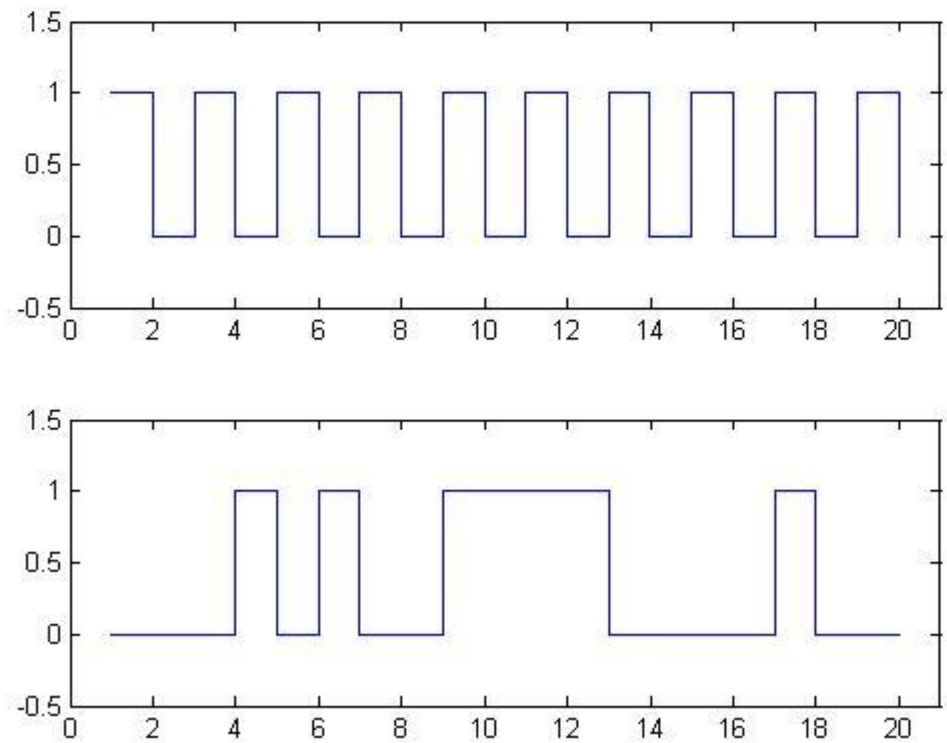
$$|\tau| \le t_a: \quad R_X(\tau) = A^2 \frac{t_a - |\tau|}{t_a} = A^2\left(1 - \frac{|\tau|}{t_a}\right)$$

$$|\tau| > t_a: \quad R_X(\tau) = 0$$

Deterministisk vs. non-deterministisk

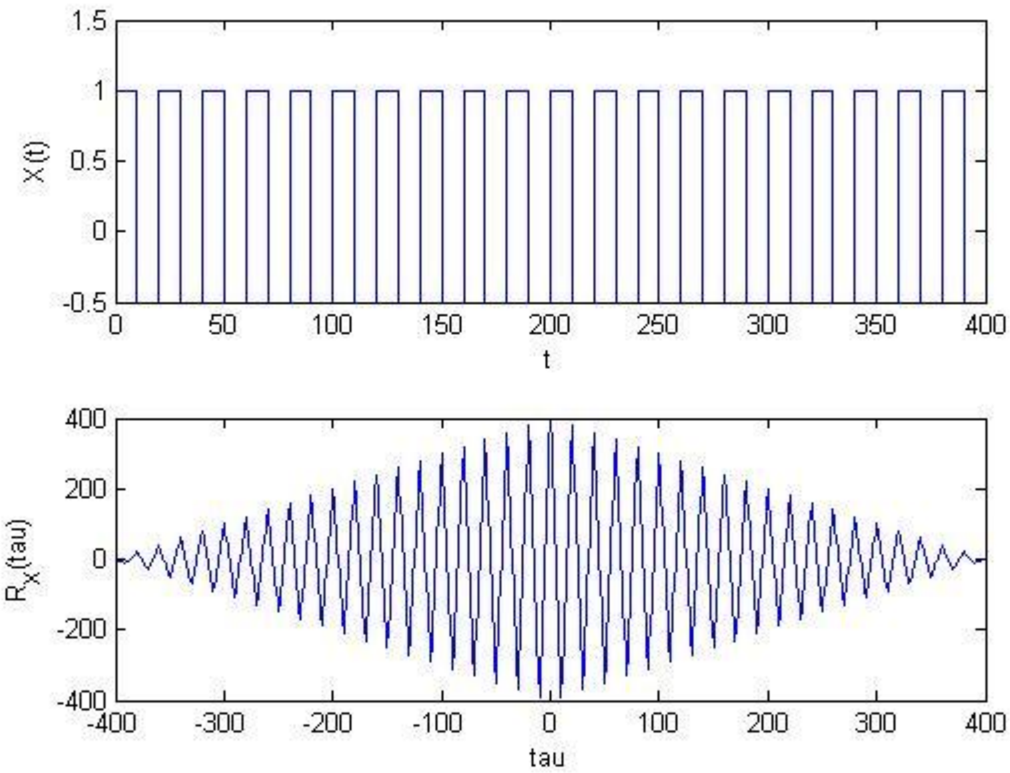
Tæthedsfunktion





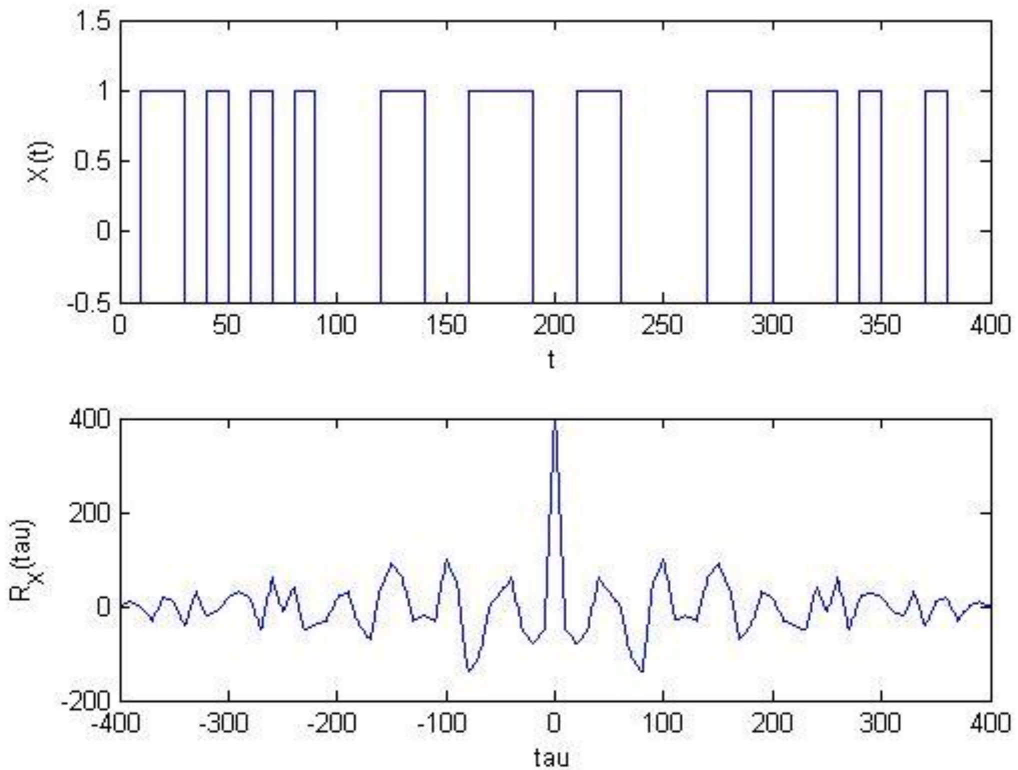
- Bemærk, at de to signaler har samme tæthedsfunktion.

Deterministisk vs. non-deterministisk



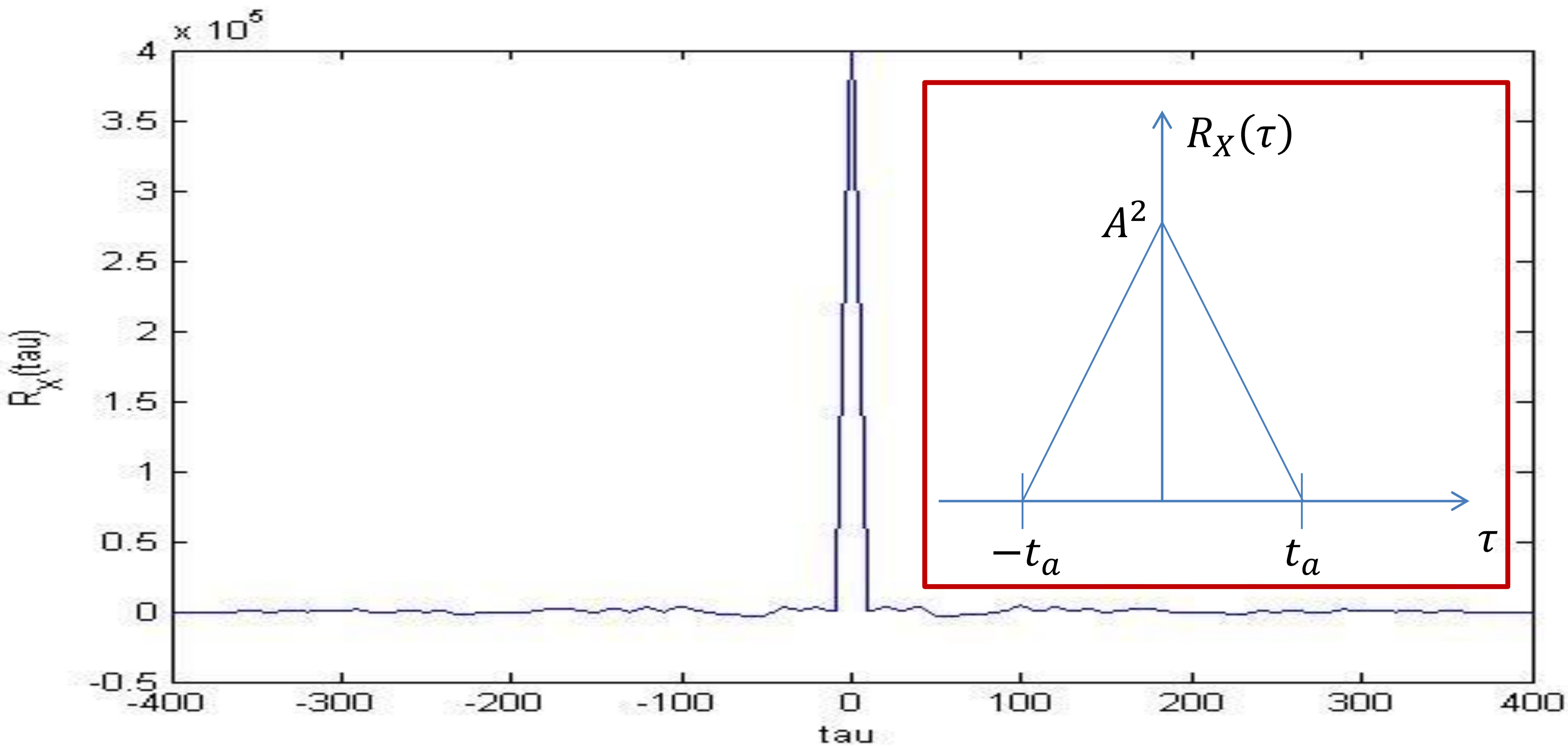

```
Rx = conv(x,fliplr(x));
```

Deterministisk vs. non-deterministisk




```
Rx = conv(x,fliplr(x));
```

Deterministisk vs. non-deterministisk



Autokorrelation midlet over 1000 simulationer

Vigtige regneregler

Egenskaber ved middelværdier og varianser

$$
\begin{aligned}
E[aX+b] &= aE[X]+b \qquad \text{middelværdien af } aX+b && (12)\\
\mathrm{Var}(aX+b) &= a^2\mathrm{Var}(X) \qquad \text{variansen af } aX+b && (13)\\
E[X+Y] &= E[X]+E[Y] \qquad \text{linearitet af middelværdi} && (14)\\
\mathrm{Var}(X+Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\cdot\mathrm{Cov}(X,Y) && (15)\\
E[X\cdot Y] &= E[X]\cdot E[Y] \qquad \text{når } X \text{ og } Y \text{ er uafhængige} && (16)\\
\mathrm{Cov}(\mathrm{X},\mathrm{Y}) &= E[(X-E[X])(Y-E[Y])] = E[X\cdot Y]-E[X]E[Y] && (17)\\
\mathrm{Corr}(\mathrm{X},\mathrm{Y}) &= \frac{\mathrm{Cov}(\mathrm{X},\mathrm{Y})}{\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}} && (18)
\end{aligned}
$$

Egenskaber ved autokorrelation

- $R_X(0) = \overline{X^2}$ (mean square)
- $R_X(\tau) = R_X(-\tau)$ (symmetri)
- $|R_X(\tau)| \leq R_X(0)$

- $\bar{X} \neq 0 \rightarrow R_X(\tau)$ indeholder en konstant komponent.
  - Se eksempel i bogen

- X indeholder en periodisk komponent $\rightarrow$ $R_X(\tau)$ indeholder en periodisk komponent.
  - Se eksempel i bogen

Måling af autokorrelation i praksis

- Husk, at for en ergodisk proces

$$R_X(\tau) = \mathcal{R}_X(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot x(t+\tau) d\tau$$

- Givet data

$$x(0), x(\Delta t), x(2\Delta t), \dots, x(N\Delta t)$$

- Så er estimatet af autokorrelationen

$$\hat{R}_X(n\Delta t) = \frac{1}{N-n+1} \sum_{k=0}^{N-n} x(k\Delta t) x\big((k+n)\Delta t\big)$$

Måling af autokorrelation i praksis

- Vi har allerede set (Matlab eksempel), at estimatet $\hat{R}_X(n\Delta t)$ er en stokastisk variabel.

- $\hat{R}_X(n\Delta t)$ er unbiased

$$E\left[\hat{R}_X(n\Delta t)\right] = R_X(n\Delta t)$$

- Om variansen af estimatet gælder

$$Var\left(\hat{R}_X(n\Delta t)\right) \leq \frac{2}{N} \sum_{k=-M}^{M} R_X^2(k\Delta t)$$

Hvor mange samples skal man bruge?

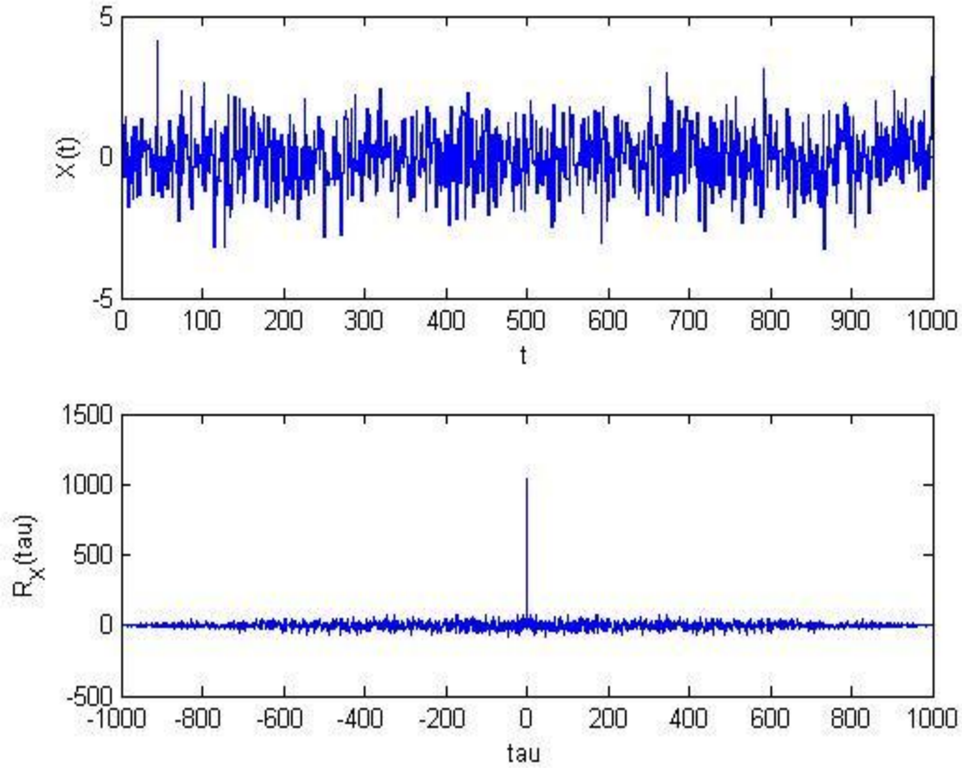
- Hvis estimatet $\hat{R}_X(n\Delta t)$ skal være godt, skal variansen af estimatet være lille.
- Eksempel:
  - Vi ser på eksemplet fra før (Figur 6-1 + 6-2).
  - Vi vælger M = 4, og så bliver $t_a = 4\Delta t$
  - Hvis RMS fejlen skal være mindre end 5%, får vi

$$Var\left(\hat{R}_X(n\Delta t)\right) \leq \frac{2}{N}\sum_{k=-M}^{M} R_X^2(k\Delta t) = \frac{2}{N}\sum_{k=-M}^{M}\left[A^2\left(1-\frac{|k|\Delta t}{4\Delta t}\right)\right]^2$$

  - Løser man mht. N, får man $N \geq 2200$

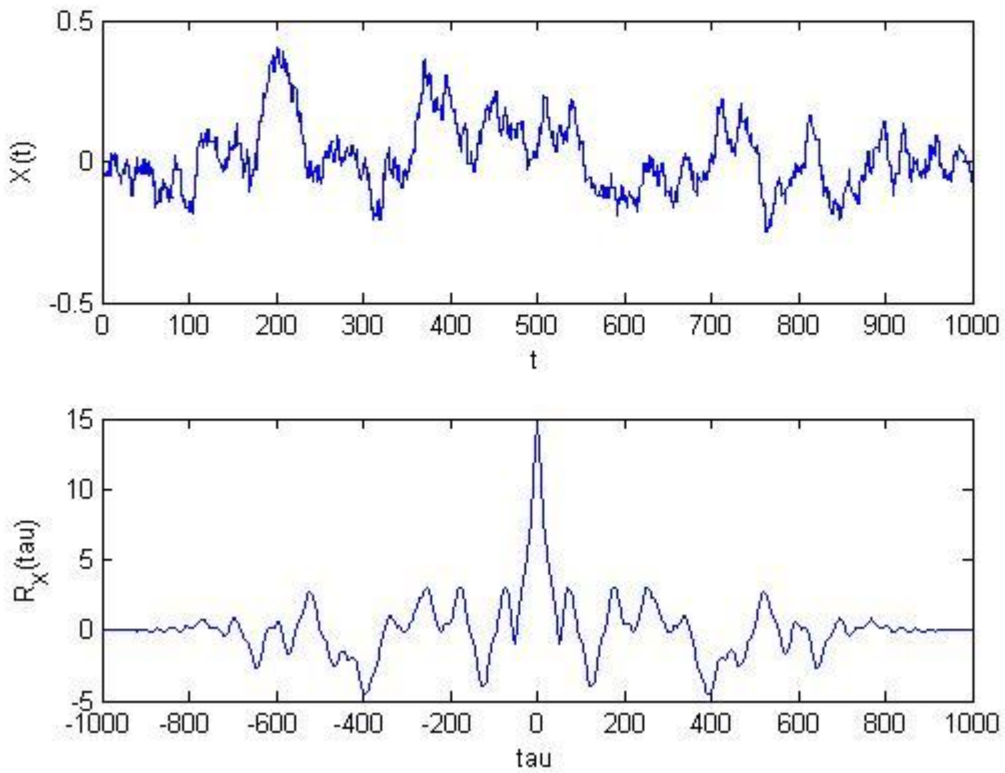
Flere eksempler

```
% Autokorrelation af hvid
t = 0:999;
tau = -999:999;
x = randn(1,1000);
Rx = conv(x,fliplr(x));
figure
subplot(2,1,1)
stairs(t,x)
xlabel('t')
ylabel('X(t)')
subplot(2,1,2)
plot(tau,Rx)
xlabel('tau')
ylabel('R_X(tau)')
```



Flere eksempler

```
% Autokorrelation af
filtreret hvid støj
t = 0:999;
tau = -999:999;
x = randn(1,1000);
h = ones(1,51)/51;
x = conv(x,h,'same');
Rx = conv(x,fliplr(x));
figure
subplot(2,1,1)
stairs(t,x)
xlabel('t')
ylabel('X(t)')
subplot(2,1,2)
plot(tau,Rx)
xlabel('tau')
ylabel('R_X(tau)')
```



Krydskorrelation

- Generelt

$$R_{XY}(t_1, t_2) = E[X(t_1) \cdot Y(t_2)] = E[X_1 \cdot Y_2]$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 \cdot y_2 \cdot f(x_1, y_2)\, dx_1 dy_2$$

- Hvis X og Y er indbyrdes stationære processer

$$R_{XY}(t_1, t_2) = R_{XY}(t_1 + T, t_2 + T) = E[X(t_1 + T) \cdot Y(t_2 + T)]$$

eller bare $R_{XY}(\tau) = E[X(t) \cdot Y(t + \tau)]$

Tidslig krydskorrelation

- Generelt

$$\mathcal{R}_{XY}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot y(t+\tau) d\tau$$

- Hvis processen er ergodisk

$$R_{XY}(\tau) = \mathcal{R}_{XY}(\tau)$$
$$R_{YX}(\tau) = \mathcal{R}_{YX}(\tau)$$