

Überblick zu Kapitel 4 des Buches **Stochastische Signaltheorie**

Es werden Zufallsgrößen mit statistischen Bindungen behandelt und anhand typischer Beispiele verdeutlicht. Nach der allgemeinen Beschreibung **zweidimensionaler Zufallsgrößen** wenden wir uns der zweidimensionalen Autokorrelationsfunktion (AKF) und der Kreuzkorrelationsfunktion (KKF) zu. Ebenso werden die dazugehörigen 2D–Spektralfunktionen angegeben.

Die AKF dient zur Beschreibung der inneren **statistischen Bindungen** eines Zufallsprozesses, während eine KKF die statistischen Abhängigkeiten zwischen zwei zwar unterschiedlichen, aber möglicherweise miteinander verkoppelten Prozessen beschreibt. Im Einzelnen werden behandelt:

- die statistische Beschreibung von *2D–Zufallsgrößen* mit Hilfe der (Verbund–)WDF,
- der Unterschied zwischen *statistischer Abhängigkeit* und *Korrelation*,
- die Klassifizierungsmerkmale *Stationarität* und *Ergodizität* stochastischer Prozesse,
- die Definitionen von *Autokorrelationsfunktion* (AKF) und *Leistungsdichtespektrum* (LDS),
- die Definitionen von *Kreuzkorrelationsfunktion* und *Kreuzleistungsdichtespektrum*, und
- die *numerische Ermittlung* all dieser Größen im zwei– und *mehrdimensionalen* Fall.

Die theoretischen Grundlagen werden auf 53 Bildschirmseiten verdeutlicht. Außerdem beinhaltet dieses Kapitel noch 16 Aufgaben und 14 Zusatzaufgaben mit insgesamt 147 Teilaufgaben sowie 88 Grafiken, fünf Lernvideos (LV) und zwei Interaktionsmodule (IM):

- **Gaußsche Zufallsgrößen ohne statistische Bindungen** (LV zu Kap. 4.1 – Dauer 2:35)
- **Gaußsche Zufallsgrößen mit statistischen Bindungen** (LV zu Kap. 4.2 – Dauer 3:15)
- **Der AWGN–Kanal (Teil 1)** (LV zu Kap. 4.5 – Dauer 6:00)
- **Der AWGN–Kanal (Teil 2)** (LV zu Kap. 4.5 – Dauer 5:15)
- **Der AWGN–Kanal (Teil 3)** (LV zu Kap. 4.5 – Dauer 6:15)
- **WDF und VTF bei Gaußschen 2D–Zufallsgrößen** (IM zu Kap. 4.2)
- **Korrelationskoeffizient und Regressionsgerade** (IM zu Kap. 4.1)

Literaturhinweise: [Böh93] – [BFS87] – [Chu78] – [Dav70] – [Hän83] – [Hän97] – [Hau03] – [Knu73] – [Knu81] – [Lük92] – [Mül79] – [Mül91] – [PP02] – [Söd88] – [Söd93] – [Win77]

Weitere Informationen zum Thema sowie Aufgaben, Simulationen und Programmierübungen finden Sie in der Anleitung zum Praktikum *Simulationsmethoden in der Nachrichtentechnik* von Prof. Günter Söder. Diese frühere LNT-Lehrveranstaltung (bis Sommersemester 2012) basiert auf den 24 DOS-Programmen des Lehrsoftwarepakets *LNTsim*.

- Kapitel 5: *Zweidimensionale Zufallsgrößen*, Programm *zwd*
- Kapitel 9: *Stochastische Prozesse*, Programm *sto*

Hinweise zum Herunterladen des Programmpakets *LNTsim* und der Versuchsanleitungen:

Lehrsoftwarepaket LNTsim (Zip–Version, mehr als 50 MB)

Praktikumsanleitung – Teil A (PDF–Version, ca. 8.5 MB, mit Kapitel 5)

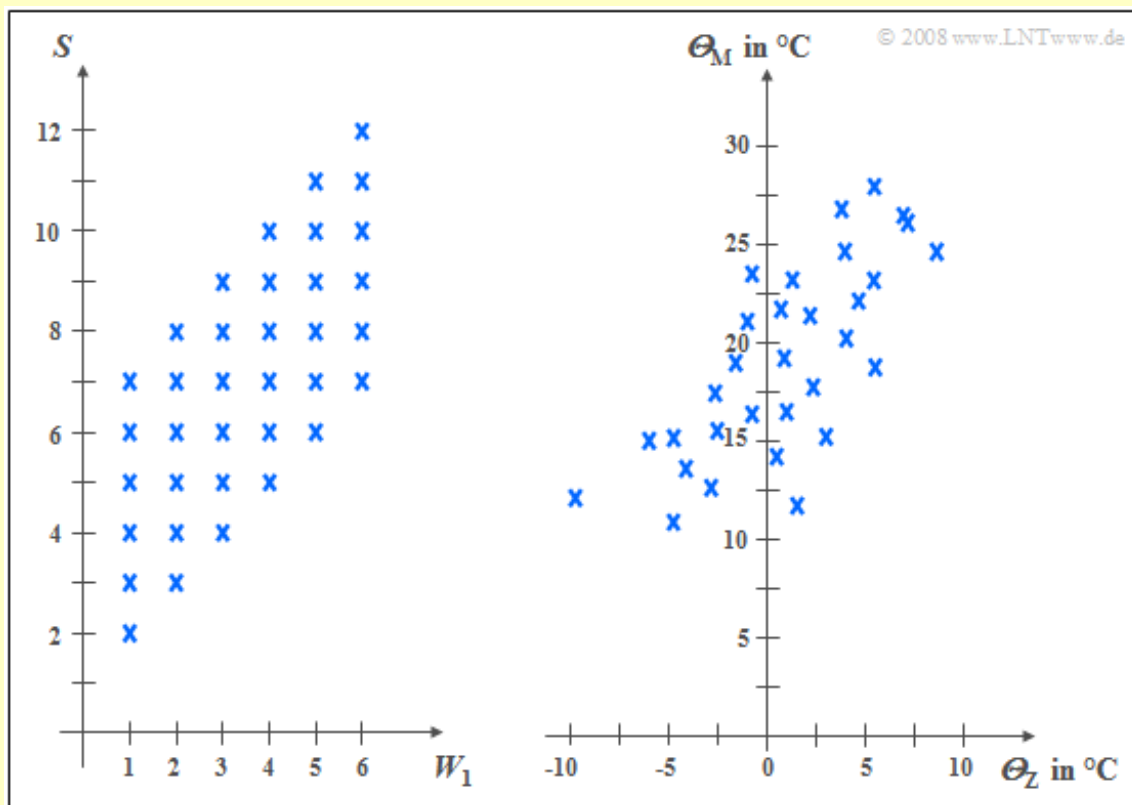
Praktikumsanleitung – Teil B (PDF–Version, ca. 9.5 MB, mit Kapitel 9)

Eigenschaften und Beispiele

Als Überleitung zu den Korrelationsfunktionen betrachten wir nun zwei Zufallsgrößen x und y , zwischen denen **statistische Abhängigkeiten** bestehen. Jede der beiden Zufallsgrößen kann für sich alleine mit den in **Kapitel 2** bzw. **Kapitel 3** eingeführten Kenngrößen beschrieben werden, je nachdem, ob es sich um eine diskrete oder um eine kontinuierliche Zufallsgröße handelt.

Zur Beschreibung der Wechselbeziehungen zwischen zwei Größen x und y ist es zweckmäßig, die beiden Komponenten zu einer **zweidimensionalen Zufallsgröße** (x, y) zusammenzufassen. Die Einzelkomponenten können Signale sein wie der Real- und Imaginärteil eines phasenmodulierten Signals. Aber es gibt auch in anderen Bereichen eine Vielzahl von 2D-Zufallsgrößen.

Beispiel: Das folgende linke Diagramm stammt von dem Zufallsexperiment *Werfen mit zwei Würfeln*. Nach rechts aufgetragen ist die Augenzahl des ersten Würfels (W_1), nach oben die Summe S beider Würfel. Die beiden Komponenten sind hier jeweils diskrete Zufallsgrößen, zwischen denen statistische Bindungen bestehen. Ist $W_1 = 1$, so kann S nur Werte zwischen 2 und 7 annehmen und zwar mit jeweils gleicher Wahrscheinlichkeit, bei $W_1 = 6$ dagegen die Werte zwischen 7 und 12.



Rechts sind die Maximaltemperaturen der 31 Tage im Mai 2002 von München (nach oben) und der Zugspitze (nach rechts) gegenübergestellt. Beide Zufallsgrößen sind wertkontinuierlich. Obwohl die Messpunkte etwa 100 km auseinander liegen und es auf der Zugspitze aufgrund der unterschiedlichen Höhenlagen (knapp 3000 gegenüber 520 Meter) im Mittel um etwa 20 Grad kälter ist als in München, erkennt man doch eine gewisse statistische Abhängigkeit zwischen den beiden Größen θ_M und θ_Z : Ist es in München warm, dann sind auch auf der Zugspitze eher angenehme Temperaturen zu erwarten. Der Zusammenhang ist aber nicht deterministisch: Der kälteste Tag im Mai 2002 war in München ein anderer als der kälteste Tag auf der Zugspitze.

Verbundwahrscheinlichkeitsdichtefunktion

Wir beschränken uns hier meist auf kontinuierliche Zufallsgrößen. Manchmal wird jedoch auch auf die Besonderheiten zweidimensionaler diskreter Zufallsgrößen genauer eingegangen.

Die meisten der bisherigen, für eindimensionale Zufallsgrößen definierten Kenngrößen können problemlos auf zweidimensionale Größen erweitert werden:

- Die Wahrscheinlichkeitsdichtefunktion der zweidimensionalen Zufallsgröße an der Stelle (x_μ, y_μ) , die man auch als **Verbundwahrscheinlichkeitsdichtefunktion** bezeichnet, ist eine Erweiterung der eindimensionalen WDF (\cap kennzeichnet die logische UND-Verknüpfung):

$$f_{xy}(x_\mu, y_\mu) = \dots = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\Pr[(x_\mu - \Delta x/2 \leq x \leq x_\mu + \Delta x/2) \cap (y_\mu - \Delta y/2 \leq y \leq y_\mu + \Delta y/2)]}{\Delta x \cdot \Delta y}.$$

Bei diskreten Zufallsgrößen ist die Definition geringfügig zu modifizieren: Bei den jeweils unteren Bereichsgrenzen ist gemäß **Kapitel 3.2** das „ \leq “-Zeichen durch das „ $<$ “-Zeichen zu ersetzen.

- Anhand dieser (Verbund)-WDF $f_{xy}(x, y)$ werden auch statistische Abhängigkeiten innerhalb der zweidimensionalen Zufallsgröße (x, y) vollständig erfasst im Gegensatz zu den beiden eindimensionalen Dichtefunktionen \Rightarrow **Randwahrscheinlichkeitsdichtefunktionen**:

$$f_x(x) = \int_{-\infty}^{+\infty} f_{xy}(x, y) dy,$$

$$f_y(y) = \int_{-\infty}^{+\infty} f_{xy}(x, y) dx.$$

Die beiden Randdichtefunktionen $f_x(x)$ und $f_y(y)$ liefern lediglich statistische Aussagen über die Einzelkomponenten x bzw. y , nicht jedoch über die Bindungen zwischen diesen.

Zweidimensionale Verteilungsfunktion

Auch die **2D-Verteilungsfunktion** ist lediglich eine sinnvolle Erweiterung der **eindimensionalen Verteilungsfunktion** (VTF):

$$F_{xy}(r_x, r_y) = \Pr [(x \leq r_x) \cap (y \leq r_y)].$$

Der Funktionalzusammenhang zwischen zweidimensionaler WDF und zweidimensionaler VTF ist wie im eindimensionalen Fall durch die Integration gegeben, aber nun in zwei Dimensionen. Bei kontinuierlichen Zufallsgrößen gilt:

$$F_{xy}(r_x, r_y) = \int_{-\infty}^{r_y} \int_{-\infty}^{r_x} f_{xy}(x, y) \, dx \, dy.$$

Umgekehrt kann auch die Wahrscheinlichkeitsdichtefunktion aus der Verteilungsfunktion durch partielle Differentiation nach r_x und r_y berechnet werden:

$$f_{xy}(x, y) = \left. \frac{d^2 F_{xy}(r_x, r_y)}{dr_x \, dr_y} \right|_{\substack{r_x=x \\ r_y=y}}.$$

Bezüglich der Verteilungsfunktion $F_{xy}(r_x, r_y)$ gelten folgende Grenzwerte:

$$F_{xy}(-\infty, -\infty) = 0,$$

$$F_{xy}(r_x, \infty) = F_x(r_x),$$

$$F_{xy}(\infty, r_y) = F_y(r_y),$$

$$F_{xy}(\infty, \infty) = 1.$$

Im Grenzfall (unendlich große r_x und r_y) ergibt sich demnach für die 2D-VTF der Wert 1. Daraus erhält man die **Normierungsbedingung** für die 2D-Wahrscheinlichkeitsdichtefunktion:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{xy}(x, y) \, dx \, dy = 1.$$

Beachten Sie den Unterschied zwischen eindimensionalen und zweidimensionalen Zufallsgrößen:

- Bei eindimensionalen Zufallsgrößen ergibt die Fläche unter der WDF stets den Wert 1.
- Bei zweidimensionalen Zufallsgrößen ist das WDF-Volumen immer gleich 1.

WDF und VTF bei statistisch unabhängigen Komponenten

Bei **statistisch unabhängigen Komponenten** x und y gilt für die Verbundwahrscheinlichkeit nach den elementaren Gesetzmäßigkeiten der Statistik, falls x und y wertkontinuierlich sind:

$$\Pr[(x_1 \leq x \leq x_2) \cap (y_1 \leq y \leq y_2)] = \Pr(x_1 \leq x \leq x_2) \cdot \Pr(y_1 \leq y \leq y_2).$$

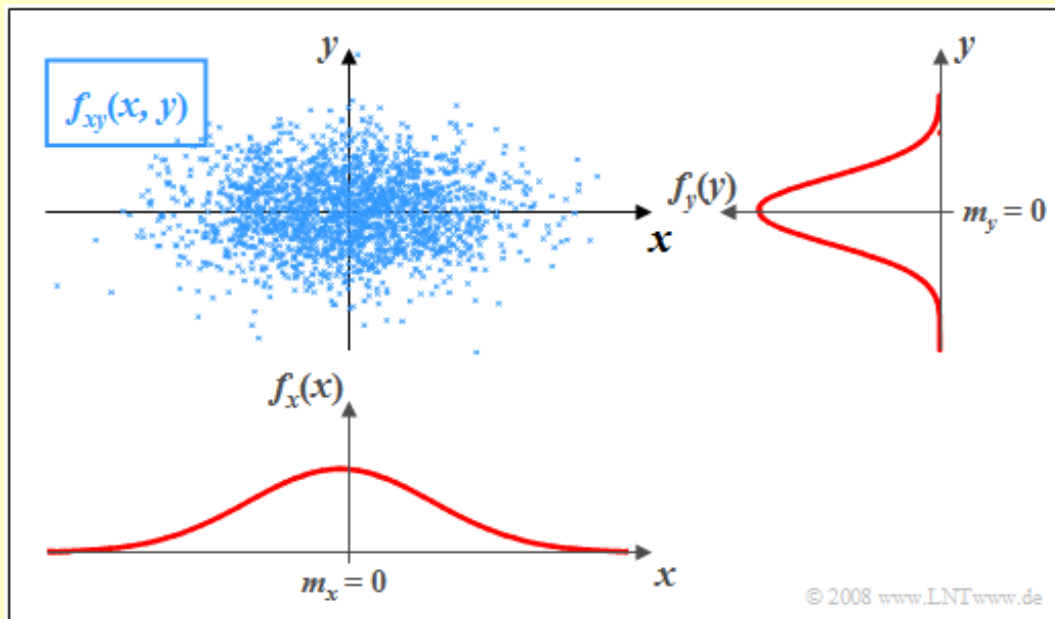
Hierfür kann bei unabhängigen Komponenten auch geschrieben werden:

$$\Pr[(x_1 \leq x \leq x_2) \cap (y_1 \leq y \leq y_2)] = \int_{x_1}^{x_2} f_x(x) dx \cdot \int_{y_1}^{y_2} f_y(y) dy.$$

Daraus folgt, dass bei statistischer Unabhängigkeit folgende Bedingung erfüllt sein muss:

$$f_{xy}(x, y) = f_x(x) \cdot f_y(y).$$

Beispiel: In der Grafik sind die Momentanwerte einer zweidimensionalen Zufallsgröße als Punkte in der (x, y) -Ebene eingetragen. Bereiche mit vielen Punkten, die dementsprechend dunkel wirken, kennzeichnen große Werte der WDF $f_{xy}(x, y)$. Dagegen besitzt die Zufallsgröße (x, y) in eher hellen Bereichen nur verhältnismäßig wenig Anteile.



Die Randwahrscheinlichkeitsdichten $f_x(x)$ und $f_y(y)$ lassen bereits erkennen, dass sowohl x als auch y gaußähnlich und mittelwertfrei sind, und dass die Zufallsgröße x eine größere Streuung als y aufweist. Sie liefern jedoch keine Informationen darüber, ob bei der Zufallsgröße (x, y) statistische Bindungen zwischen den beiden Komponenten bestehen oder nicht.

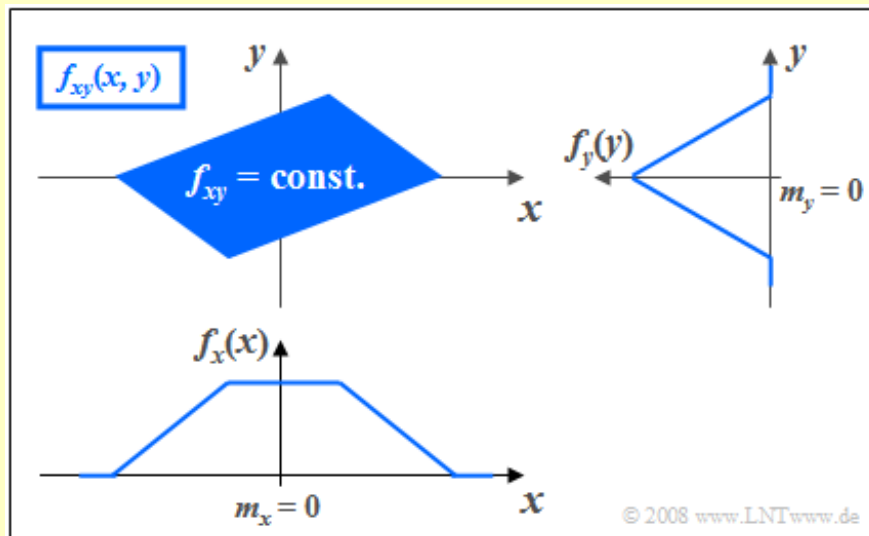
Anhand der 2D-WDF ist zu erkennen, dass es hier keine statistischen Bindungen zwischen den Komponenten gibt. Bei statistischer Unabhängigkeit liefert jeder Schnitt durch $f_{xy}(x, y)$ parallel zur y -Achse eine Funktion, die formgleich mit der Randwahrscheinlichkeitsdichtefunktion $f_y(y)$ ist. Ebenso sind alle Schnitte parallel zur x -Achse formgleich mit $f_x(x)$.

Diese Tatsache ist gleichbedeutend mit der Aussage, dass die 2D-WDF $f_{xy}(x, y)$ als Produkt der beiden Randwahrscheinlichkeitsdichten dargestellt werden kann.

WDF und VTF bei statistisch abhängigen Komponenten

Bestehen statistische Bindungen zwischen den Komponenten, so liefern unterschiedliche Schnitte parallel zur x - bzw. y -Achse jeweils unterschiedliche, nicht formgleiche Funktionen. In diesem Fall lässt sich die Verbund-Wahrscheinlichkeitsdichtefunktion natürlich auch nicht als Produkt der beiden (eindimensionalen) Randwahrscheinlichkeitsdichten beschreiben.

Beispiel: Die Grafik zeigt die Momentanwerte einer zweidimensionalen Zufallsgröße in der (x, y) -Ebene, wobei zwischen x und y nun statistische Bindungen bestehen. Die 2D-Zufallsgröße nimmt in dem blau eingezeichneten Parallelogramm alle Werte mit gleicher Wahrscheinlichkeit an, außerhalb sind keine Werte möglich.



Die Integration über die 2D-WDF $f_{xy}(x, y)$ parallel zu der x -Achse führt zur dreieckförmigen Randdichte $f_y(y)$, die Integration parallel zur y -Achse zur trapezförmigen WDF $f_x(x)$.

Aus der zweidimensionalen Wahrscheinlichkeitsdichtefunktion $f_{xy}(x, y)$ ist bereits zu erahnen, dass für jeden x -Wert im statistischen Mittel ein anderer y -Wert zu erwarten ist. Daran erkennt man ebenfalls, dass hier die Komponenten x und y statistisch voneinander abhängen.

Erwartungswerte zweidimensionaler Zufallsgrößen (1)

Ein Sonderfall der statistischen Abhängigkeit ist die **Korrelation**. Darunter versteht man eine lineare Abhängigkeit zwischen den Einzelkomponenten x und y .

- Korrelierte Zufallsgrößen sind damit stets auch statistisch abhängig.
- Aber nicht jede statistische Abhängigkeit beschreibt gleichzeitig eine Korrelation.

Zur quantitativen Erfassung der Korrelation verwendet man verschiedene Erwartungswerte der 2D-Zufallsgröße (x, y) , die analog zum eindimensionalen Fall nach **Kapitel 2.2** bzw. **Kapitel 3.3** definiert sind:

- Für die (nichtzentrierten) **Momente** gilt die Beziehung:

$$m_{kl} = E[x^k \cdot y^l] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^k \cdot y^l \cdot f_{xy}(x, y) dx dy.$$

Somit sind die beiden linearen Mittelwerte $m_x = m_{10}$ und $m_y = m_{01}$.

- Die auf m_x bzw. m_y bezogenen **Zentralmomente** lauten:

$$\mu_{kl} = E[(x - m_x)^k \cdot (y - m_y)^l].$$

In dieser allgemein gültigen Definitionsgleichung sind die Varianzen σ_x^2 und σ_y^2 der zwei Einzelkomponenten durch μ_{20} bzw. μ_{02} mit enthalten.

- Besondere Bedeutung besitzt die sogenannte **Kovarianz** ($k = l = 1$), die ein Maß für die lineare statistische Abhängigkeit zwischen den Zufallsgrößen x und y ist:

$$\mu_{11} = E[(x - m_x) \cdot (y - m_y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_x)(y - m_y) \cdot f_{xy}(x, y) dx dy.$$

Im Folgenden bezeichnen wir die Kovarianz μ_{11} teilweise auch mit μ_{xy} , falls sich die Kovarianz auf die Zufallsgrößen x und y bezieht. Die Kovarianz hängt wie folgt mit dem nichtzentrierten Moment $m_{11} = m_{xy} = E[x \cdot y]$ zusammen:

$$\mu_{xy} = m_{xy} - m_x \cdot m_y.$$

Diese Gleichung ist für die numerische Auswertung von enormen Vorteil, da m_{xy} , m_x und m_y aus den Folgen $\langle x_v \rangle$ und $\langle y_v \rangle$ direkt - also in einem Durchlauf - gefunden werden können. Würde man dagegen die Kovarianz μ_{xy} entsprechend der oberen Definitionsgleichung direkt berechnen, so müsste man in einem ersten Durchlauf die Mittelwerte m_x und m_y ermitteln und dann in einem zweiten Durchlauf den Erwartungswert $E[(x - m_x) \cdot (y - m_y)]$.

Erwartungswerte zweidimensionaler Zufallsgrößen (2)

Beispiel: Nachfolgend sehen Sie die jeweils ersten Elemente zweier Zufallsfolgen $\langle x_v \rangle$ und $\langle y_v \rangle$. In der letzten Zeile sind die jeweiligen Produkte $x_v \cdot y_v$ angegeben.

x_v	0.21	0.48	0.57	0.93	0.45	0.17	0.28	0.35	0.68	0.88
y_v	0.13	0.94	1.21	2.29	0.85	0.01	0.34	0.55	1.54	2.14
$x_v \cdot y_v$	0.0273	0.4512	0.6897	2.1297	0.3825	0.0017	0.0952	0.1925	1.0472	1.8832

© 2008 www.LNTwww.de

- Durch Mittelung über die jeweils zehn Folgeelemente erhält man $m_x = 0.5$, $m_y = 1$ und $m_{xy} = 0.69$. Daraus ergibt sich die Kovarianz zu $\mu_{xy} = 0.69 - 0.5 \cdot 1 = 0.19$.
- Ohne Kenntnis der Gleichung $\mu_{xy} = m_{xy} - m_x \cdot m_y$ hätte man zunächst im ersten Durchlauf die Mittelwerte m_x und m_y ermitteln müssen, um im zweiten Durchlauf die Kovarianz μ_{xy} als Erwartungswert des Produkts der mittelwertfreien Größen bestimmen zu können.

Korrelationskoeffizient

Bei statistischer Unabhängigkeit der beiden Komponenten x und y ist die Kovarianz $\mu_{xy} = 0$. Das Ergebnis „ $\mu_{xy} = 0$ “ erhält man auch bei statistisch abhängigen Komponenten x und y , wenn diese unkorreliert \Rightarrow **linear** unabhängig sind.

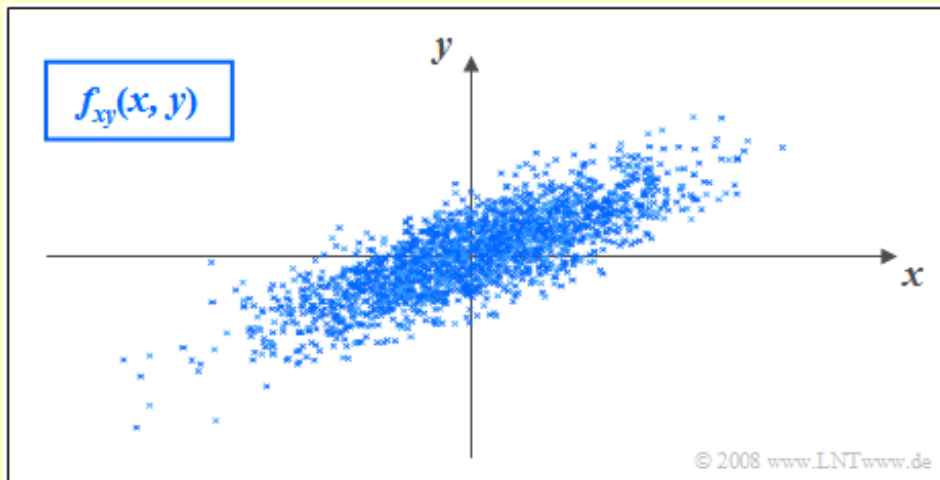
Sind dagegen x und y voll korreliert (z. B.: $y = K \cdot x$), so ergibt sich bei positivem Wert von K für die Kovarianz: $\mu_{xy} = \sigma_x \cdot \sigma_y$. Deshalb verwendet man als Beschreibungsgröße häufig anstelle der Kovarianz den **Korrelationskoeffizienten**:

$$\rho_{xy} = \frac{\mu_{xy}}{\sigma_x \cdot \sigma_y}.$$

Dieser weist folgende Eigenschaften auf:

- Aufgrund der Normierung gilt stets $-1 \leq \rho_{xy} \leq +1$.
- Sind die beiden Zufallsgrößen x und y unkorreliert, so ist $\rho_{xy} = 0$.
- Bei strenger linearer Abhängigkeit (x und y sind zueinander proportional) ist $\rho_{xy} = \pm 1$.

Beispiel: Die Zufallsgrößen x und y sind wie im Beispiel auf **Seite 4** dieses Kapitels gaußförmig verteilt, wobei die Streuungen unterschiedlich sind $\Rightarrow \sigma_y < \sigma_x$.



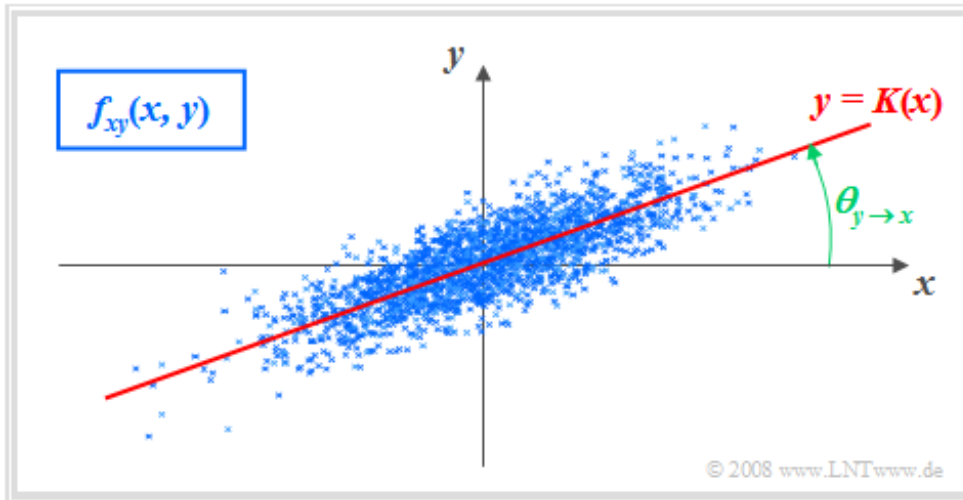
Im Unterschied zum ersten Beispiel sind aber nun die Zufallsgrößen x und y (positiv) korreliert. Der Korrelationskoeffizient ist dabei $\rho_{xy} = 0.8$.

Fassen wir zusammen:

- Ein positiver Korrelationskoeffizient bedeutet, dass bei größerem x -Wert im statistischen Mittel auch y größer ist als bei kleinerem x .
- Dagegen drückt ein negativer Korrelationskoeffizient aus, dass y mit steigendem x im Mittel kleiner wird.

Korrelationsgerade

Man kann nun in die (x, y) -Ebene eine Gerade durch den „Mittelpunkt“ (m_x, m_y) einzeichnen. Diese Gerade $y = K(x)$ bezeichnet man als **Korrelationsgerade** (oder als Regressionsgerade).



Die Korrelationsgerade besitzt folgende Eigenschaften:

- Die mittlere quadratische Abweichung von dieser Geraden – in y -Richtung betrachtet und über alle N Punkte gemittelt – ist minimal:

$$\overline{\varepsilon_y^2} = \frac{1}{N} \cdot \sum_{\nu=1}^N [y_\nu - K(x_\nu)]^2 = \text{Minimum.}$$

- Die Korrelationsgerade kann man als eine Art „statistische Symmetrieachse“ interpretieren. Die Geradengleichung lautet:

$$y = K(x) = \frac{\sigma_y}{\sigma_x} \cdot \rho_{xy} \cdot (x - m_x) + m_y.$$

- Der Winkel, den die Korrelationsgerade zur x -Achse einnimmt, beträgt:

$$\theta_{y \rightarrow x} = \arctan\left(\frac{\sigma_y}{\sigma_x} \cdot \rho_{xy}\right).$$

Durch diese Nomenklatur soll deutlich gemacht werden, dass es sich hier um die Regression von y auf x handelt. Die Regression in Gegenrichtung – also von x auf y – bedeutet dagegen die Minimierung der mittleren quadratischen Abweichung in x -Richtung.

Wie die folgende Animation verdeutlicht, ergibt sich dafür im Allgemeinen eine andere Gerade.

Korrelationskoeffizient und Regressionsgerade

Wahrscheinlichkeitsdichte- und Verteilungsfunktion (1)

Alle bisherigen Aussagen von Kapitel 4 gelten allgemein. Für den Sonderfall **Gaußscher Zufallsgrößen** – der Name geht auf den Wissenschaftler **Carl Friedrich Gauß** zurück – können wir weiterhin vermerken:

- Die Verbundwahrscheinlichkeitsdichtefunktion einer Gaußschen 2D-Zufallsgröße (x, y) mit den Mittelwerten $m_x = 0$ und $m_y = 0$ sowie dem Korrelationskoeffizienten ρ_{xy} lautet:

$$f_{xy}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} \cdot \exp\left[-\frac{1}{2(1-\rho_{xy}^2)} \cdot \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - 2\rho_{xy} \cdot \frac{x \cdot y}{\sigma_x \cdot \sigma_y}\right)\right].$$

- Ersetzt man in dieser Gleichung x durch $(x - m_x)$ sowie y durch $(y - m_y)$, so ergibt sich die allgemeinere WDF einer zweidimensionalen Gaußschen Zufallsgröße mit Mittelwert.
- Die beiden Randwahrscheinlichkeitsdichtefunktionen $f_x(x)$ und $f_y(y)$ sind in diesem Fall ebenfalls gaußförmig und weisen die Streuungen σ_x bzw. σ_y auf.
- Bei unkorrelierten Komponenten x und y muss in obiger Gleichung $\rho_{xy} = 0$ eingesetzt werden, und man erhält dann das Ergebnis:

$$f_{xy}(x, y) = \frac{1}{\sqrt{2\pi} \cdot \sigma_x} \cdot e^{-x^2/(2\sigma_x^2)} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma_y} \cdot e^{-y^2/(2\sigma_y^2)} = f_x(x) \cdot f_y(y).$$

Resümee: Im Sonderfall einer **2D-Zufallsgröße mit Gaußscher WDF** $f_{xy}(x, y)$ folgt aus der *Unkorreliertheit* auch direkt die *statistische Unabhängigkeit*:

$$f_{xy}(x, y) = f_x(x) \cdot f_y(y).$$

Bei keiner anderen WDF kann aus der *Unkorreliertheit* auf die *statistische Unabhängigkeit* geschlossen werden. Man kann aber stets \Rightarrow für jede beliebige 2D-WDF $f_{xy}(x, y)$ von der *statistischen Unabhängigkeit* auf die *Unkorreliertheit* schließen, weil:

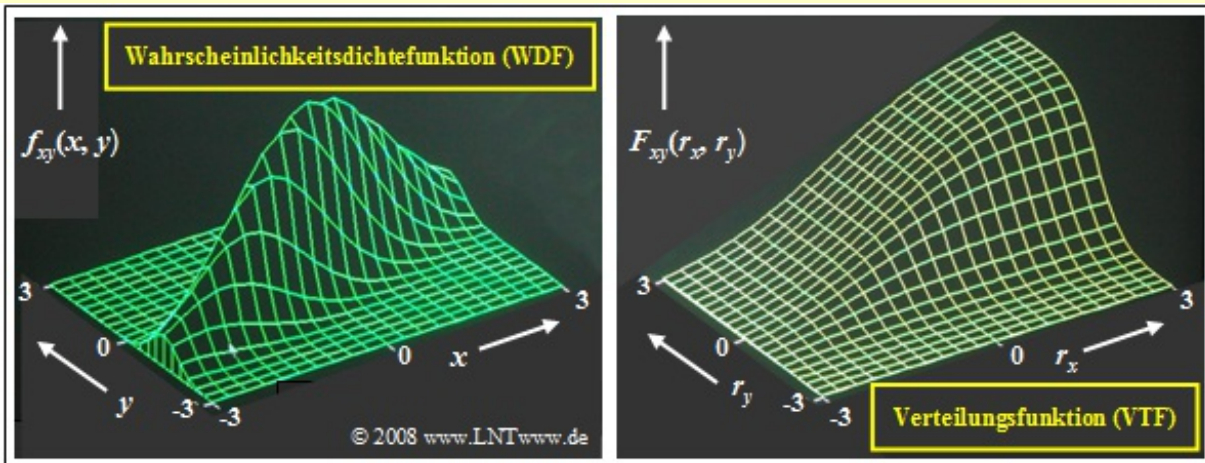
- Sind zwei Zufallsgrößen x und y völlig voneinander (statistisch) unabhängig, so gibt es zwischen ihnen natürlich auch keine *linearen* Abhängigkeiten.

Wahrscheinlichkeitsdichte- und Verteilungsfunktion (2)

Beispiel: Das Bild zeigt

- die Wahrscheinlichkeitsdichtefunktion (links) und
- Verteilungsfunktion (rechts)

einer zweidimensionalen Gaußschen Zufallsgröße (x, y) mit relativ starker positiver Korrelation der Einzelkomponenten: $\rho_{xy} = 0.8$. Wie bei den bisherigen Bildern in diesem Kapitel ist die 2D-Zufallsgröße in x -Richtung weiter ausgedehnt als in y -Richtung: $\sigma_x = 2 \cdot \sigma_y$.



Diese Darstellungen können wie folgt interpretiert werden:

- Die WDF ist vergleichbar mit einem Bergkamm, der sich von links unten nach rechts oben erstreckt.
- Das Maximum liegt bei $m_x = 0$ und $m_y = 0$. Das bedeutet, dass die 2D-Zufallsgröße mittelwertfrei ist.
- Die zweidimensionale VTF als das Integral in zwei Richtungen über die WDF steigt von links unten nach rechts oben von 0 auf 1 kontinuierlich an.

Das nachfolgende Interaktionsmodul erlaubt die Darstellung der zweidimensionalen WDF und der zweidimensionalen VTF für beliebige Werte von σ_x , σ_y und ρ_{xy} :

WDF/VTF bei 2D-Gaußgrößen

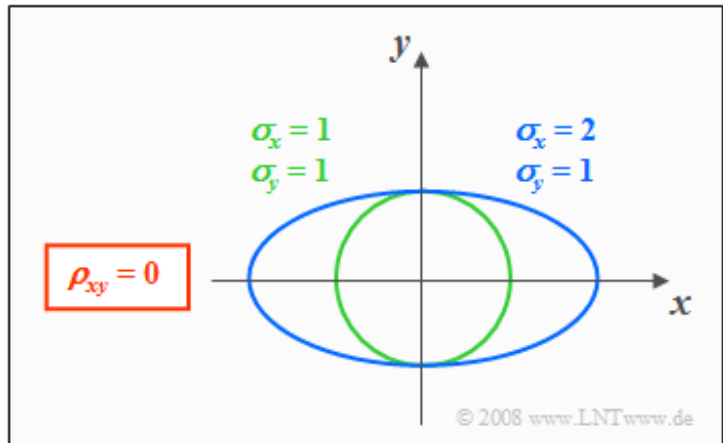
Höhenlinien bei unkorrelierten Zufallsgrößen

Aus der Bedingungsgleichung $f_{xy}(x, y) = \text{const.}$ können die **Höhenlinien der WDF** berechnet werden. Sind die Komponenten x und y unkorreliert ($\rho_{xy} = 0$), so erhält man:

$$\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} = \text{const.}$$

Die Höhenlinien beschreiben in diesem Fall folgende Figuren:

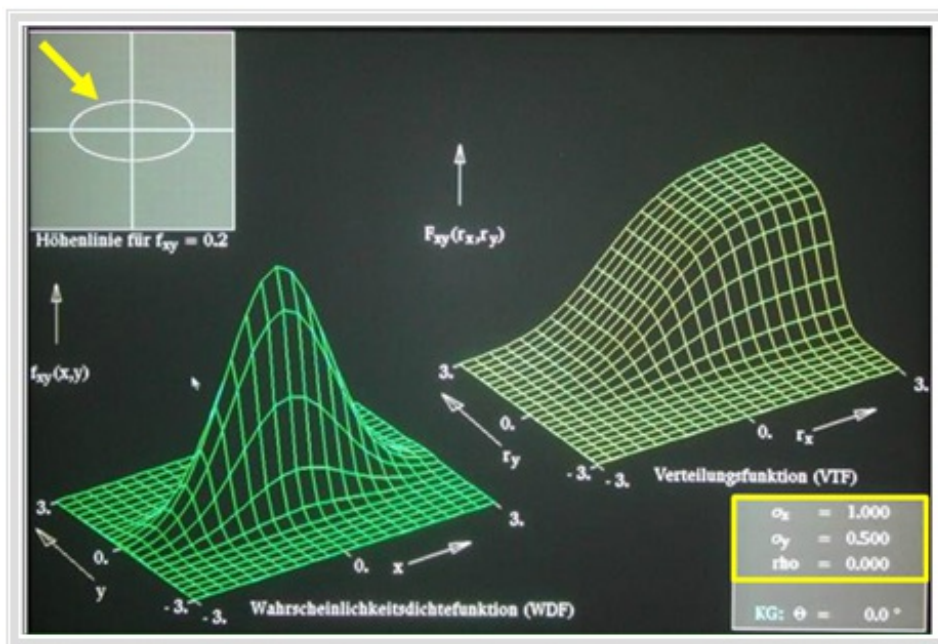
- **Kreise** (falls $\sigma_x = \sigma_y$, grüne Kurve),
oder
- **Ellipsen** (für $\sigma_x \neq \sigma_y$, blaue Kurve) in Ausrichtung der beiden Achsen.



Weitere Informationen zu dieser Thematik mit Signalbeispielen bietet das folgende Lernvideo:

Gaußsche Zufallsgrößen ohne statistische Bindungen (Dauer 2:35).

Sie sehen hier einen Bildschirmabzug dieses Multimedia-Moduls.

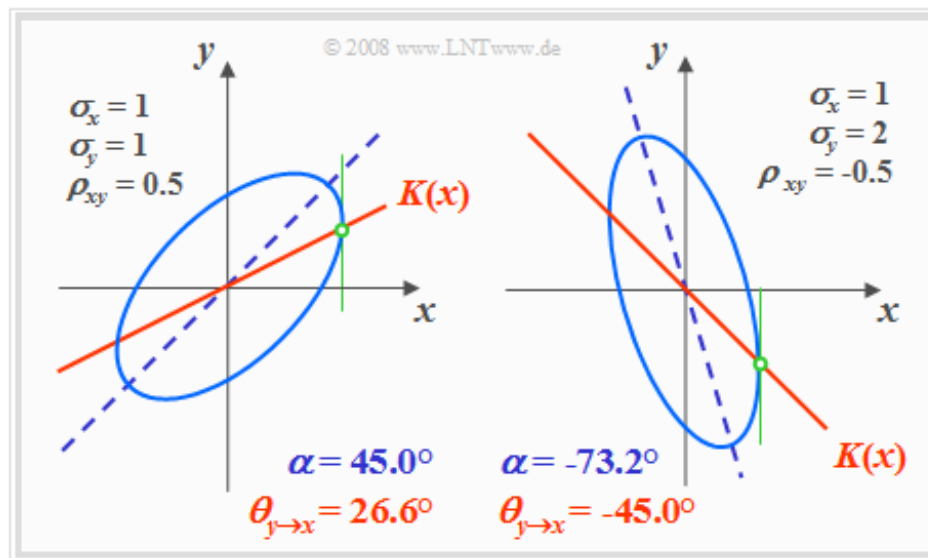


Höhenlinien bei korrelierten Zufallsgrößen

Bei korrelierten Komponenten ($\rho_{xy} \neq 0$) sind die Höhenlinien der WDF stets elliptisch, also auch für den Sonderfall $\sigma_x = \sigma_y$. Hier lautet die Bedingungsgleichung $f_{xy}(x, y) = \text{const.}$:

$$\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - 2 \cdot \rho_{xy} \cdot \frac{x \cdot y}{\sigma_x \cdot \sigma_y} = \text{const.}$$

Das folgende Bild zeigt in hellerem Blau zwei Höhenlinien für unterschiedliche Parametersätze, jeweils mit $\rho_{xy} \neq 0$. Die Ellipsenhauptachse ist dunkelblau gestrichelt, und die Korrelationsachse $K(x)$ durchgehend rot eingezeichnet.



Anhand dieses Bildes sind folgende Aussagen möglich:

- Die Ellipsenform hängt außer vom Korrelationskoeffizienten ρ_{xy} auch vom Verhältnis der beiden Streuungen σ_x und σ_y ab.
- Auch der Neigungswinkel α der Ellipsenhauptachse (gestrichelte Gerade) gegenüber der x -Achse hängt von diesen drei Parametern ab:

$$\alpha = \frac{1}{2} \cdot \arctan\left(2 \cdot \rho_{xy} \cdot \frac{\sigma_x \cdot \sigma_y}{\sigma_x^2 - \sigma_y^2}\right).$$

- Die Korrelationsgerade $y = K(x)$ einer Gaußschen 2D-Zufallsgröße liegt stets unterhalb der Ellipsenhauptachse.
- $K(x)$ kann auch aus dem Schnittpunkt der Höhenlinien und ihrer vertikalen Tangenten geometrisch konstruiert werden, wie in den obigen Skizzen in grüner Farbe angedeutet ist.

Die folgenden Lernvideos beschreiben die Eigenschaften Gaußscher Zufallsgrößen:

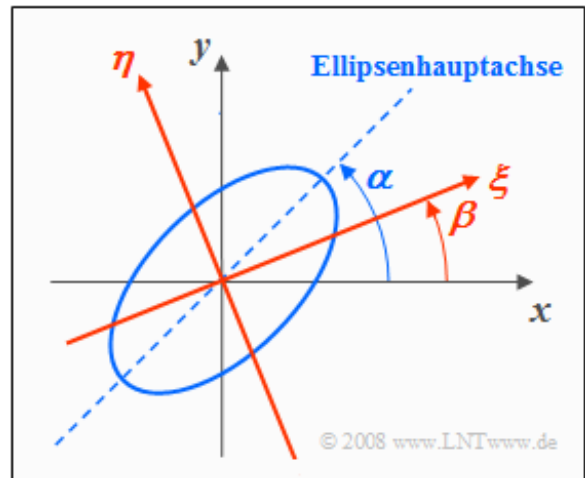
Gaußsche Zufallsgrößen ohne statistische Bindungen (Dauer 2:35),

Gaußsche Zufallsgrößen mit statistischen Bindungen (Dauer 3:05).

Drehung des Koordinatensystems (1)

Bei manchen Aufgabenstellungen ist es vorteilhaft, das Koordinatensystem zu drehen, wie in der nachfolgenden Grafik angedeutet:

- Das (ξ, η) -Koordinatensystem ist gegenüber dem ursprünglichen (x, y) -System um den Winkel β gedreht.
- Dagegen bezeichnet α den Winkel zwischen der Ellipsenhauptachse und der x -Achse.



Zwischen den Koordinaten der beiden Bezugssysteme bestehen folgende Zusammenhänge:

$$\begin{aligned} \xi &= \cos(\beta) \cdot x + \sin(\beta) \cdot y & \text{bzw.} & \quad x = \cos(\beta) \cdot \xi - \sin(\beta) \cdot \eta, \\ \eta &= -\sin(\beta) \cdot x + \cos(\beta) \cdot y & \text{bzw.} & \quad y = \sin(\beta) \cdot \xi + \cos(\beta) \cdot \eta. \end{aligned}$$

Ist (x, y) eine Gaußsche 2D-Zufallsgröße, so ist die neue Zufallsgröße (ξ, η) ebenfalls gaußverteilt.

Setzt man die obigen Gleichungen in die 2D-WDF $f_{xy}(x, y)$ ein und vergleicht die Koeffizienten, so erhält man folgende Bestimmungsgleichungen für σ_x, σ_y und ρ_{xy} bzw. für σ_ξ, σ_η und $\rho_{\xi\eta}$:

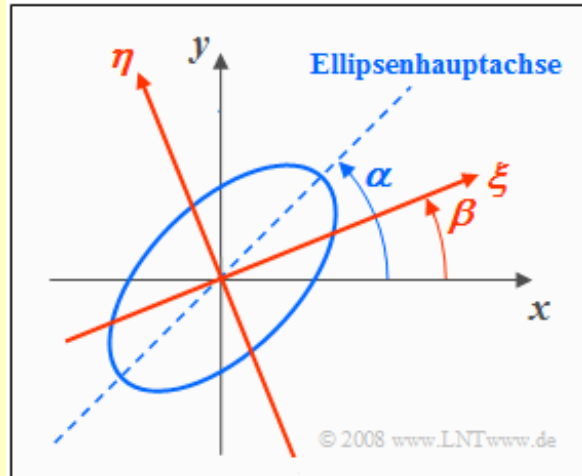
$$\begin{aligned} \frac{1}{(1 - \rho_{\xi\eta}^2) \cdot \sigma_\xi^2} &= \frac{1}{(1 - \rho_{xy}^2)} \left[\frac{\cos^2(\beta)}{\sigma_x^2} + \frac{\sin^2(\beta)}{\sigma_y^2} - 2\rho_{xy} \cdot \frac{\sin(\beta) \cdot \cos(\beta)}{\sigma_x \cdot \sigma_y} \right], \\ \frac{1}{(1 - \rho_{\xi\eta}^2) \cdot \sigma_\eta^2} &= \frac{1}{(1 - \rho_{xy}^2)} \left[\frac{\sin^2(\beta)}{\sigma_x^2} + \frac{\cos^2(\beta)}{\sigma_y^2} + 2\rho_{xy} \cdot \frac{\sin(\beta) \cdot \cos(\beta)}{\sigma_x \cdot \sigma_y} \right], \\ \frac{\rho_{\xi\eta}}{(1 - \rho_{\xi\eta}^2) \cdot \sigma_\xi \cdot \sigma_\eta} &= \\ &= \frac{1}{(1 - \rho_{xy}^2)} \left[\frac{\sin(\beta) \cdot \cos(\beta)}{\sigma_x^2} - \frac{\sin(\beta) \cdot \cos(\beta)}{\sigma_y^2} + \frac{\rho_{xy}}{\sigma_x \cdot \sigma_y} \cdot (\cos^2(\beta) - \sin^2(\beta)) \right]. \end{aligned}$$

Mit diesen drei Gleichungen können die jeweils drei Parameter der beiden Koordinatensysteme direkt umgerechnet werden, was allerdings nur in Sonderfällen ohne erheblichen Rechenaufwand möglich ist. Ein solches Beispiel folgt auf der nächsten Seite.

Drehung des Koordinatensystems (2)

Beispiel: Wir betrachten eine Gaußsche 2D-WDF mit $\sigma_x = \sigma_y = 1$ und $\rho_{xy} = 0.5$. In diesem Fall beträgt der Winkel α der Ellipsenhauptachse 45° .

- Dreht man das Koordinatensystem ebenfalls um ($\beta =$) 45° , so ergibt sich wegen $\sigma_x = \sigma_y$ und wegen $\sin(\beta) = \cos(\beta) = 0.5^{1/2}$ für den neuen Korrelationskoeffizienten $\rho_{\xi\eta} = 0$, was einsichtig ist.
- Die beiden Streuungen – bezogen auf das neue Koordinatensystem – ergeben sich dann entsprechend den beiden ersten oberen Gleichungen zu $\sigma_\xi = 1.5^{1/2}$ und $\sigma_\eta = 0.5^{1/2}$.



Dem obigen Bild ist allerdings nicht $\beta = \alpha$ zugrundegelegt, sondern $\beta = \alpha/2$. Dann lautet das Gleichungssystem mit $\sigma_x = \sigma_y = 1, \rho_{xy} = 0.5, \alpha = 45^\circ, \sin(\beta) \cdot \cos(\beta) = \sin(2\beta)/2 = \sin(\alpha)/2$ und $\cos^2(\beta) - \sin^2(\beta) = \cos(2\beta) = \cos(\alpha)$:

$$(I) \quad \frac{1}{(1 - \rho_{\xi\eta}^2) \cdot \sigma_\xi^2} = \frac{4}{3} \left[1 - \frac{1}{2} \cdot \sin(\alpha) \right] = 0.862,$$

$$(II) \quad \frac{1}{(1 - \rho_{\xi\eta}^2) \cdot \sigma_\eta^2} = \frac{4}{3} \left[1 + \frac{1}{2} \cdot \sin(\alpha) \right] = 1.805, \quad \frac{(I)}{(II)} : \frac{\sigma_\eta}{\sigma_\xi} = \sqrt{\frac{0.862}{1.805}} = 0.691,$$

$$(III) \quad \frac{\rho_{\xi\eta}}{(1 - \rho_{\xi\eta}^2) \cdot \sigma_\xi \cdot \sigma_\eta} = \frac{\rho_{\xi\eta}}{(1 - \rho_{\xi\eta}^2) \cdot \sigma_\xi^2 \cdot 0.691} = \frac{2}{3} \cdot \cos(\alpha) = 0.471.$$

Dividiert man nun die Gleichung (III) durch die Gleichung (I), so erhält man:

$$\frac{\rho_{\xi\eta}}{0.691} = \frac{0.471}{0.862} \quad \Rightarrow \quad \rho_{\xi\eta} = 0.378.$$

Die beiden weiteren Parameter ergeben sich zu $\sigma_\xi \approx 1$ und $\sigma_\eta \approx 0.7$.

Voraussetzungen und Mittelwerte

In diesem Kapitel 4.3 gehen wir von den folgenden Annahmen aus:

- Die Zufallsgrößen u und v seien jeweils mittelwertfrei $\Rightarrow m_u = m_v = 0$ und zudem statistisch unabhängig voneinander $\Rightarrow \rho_{uv} = 0$.
- Die beiden Zufallsgrößen u und v besitzen jeweils gleiche Streuung σ . Über die Art der Verteilung wird keine Aussage getroffen.
- Die beiden Zufallsgrößen x und y seien Linearkombinationen von u und v , wobei gilt:

$$x = A \cdot u + B \cdot v + C,$$

$$y = D \cdot u + E \cdot v + F.$$

Für die (linearen) Mittelwerte der neuen Zufallsgrößen x und y erhält man nach den allgemeinen Rechenregeln für Erwartungswerte:

$$m_x = A \cdot m_u + B \cdot m_v + C = C,$$

$$m_y = D \cdot m_u + E \cdot m_v + F = F.$$

Die Koeffizienten C und F geben somit lediglich die Mittelwerte von x und y an. Beide werden auf den folgenden Seiten stets zu 0 gesetzt.

Resultierender Korrelationskoeffizient

Betrachten wir nun die **Varianzen** nach den Linearkombinationen. Für die Zufallsgröße x gilt unabhängig vom Parameter C :

$$\sigma_x^2 = E[x^2] = A^2 \cdot E[u^2] + B^2 \cdot E[v^2] + 2 \cdot A \cdot B \cdot E[u \cdot v].$$

Die Erwartungswerte von u^2 und v^2 sind definitionsgemäß jeweils gleich σ^2 . Da u und v als statistisch unabhängig vorausgesetzt werden, kann man für den Erwartungswert des Produktes auch schreiben:

$$E[u \cdot v] = E[u] \cdot E[v] = m_u \cdot m_v = 0.$$

Damit erhält man für die Varianzen der durch Linearkombinationen gebildeten Zufallsgrößen:

$$\sigma_x^2 = (A^2 + B^2) \cdot \sigma^2,$$

$$\sigma_y^2 = (D^2 + E^2) \cdot \sigma^2.$$

Die **Kovarianz** μ_{xy} ist bei mittelwertfreien Zufallsgrößen x und $y \Rightarrow C = F = 0$ identisch mit dem gemeinsamen Moment m_{xy} :

$$\mu_{xy} = m_{xy} = E[x \cdot y] = E[(A \cdot u + B \cdot v)(D \cdot u + E \cdot v)].$$

Beachten Sie hierbei, dass $E[\]$ einen Erwartungswert bezeichnet, während E eine Variable beschreibt. Nach Auswertung dieser Gleichung in analoger Weise zu oben folgt daraus:

$$\begin{aligned} \mu_{xy} &= (A \cdot D + B \cdot E) \cdot \sigma^2 \\ \Rightarrow \rho_{xy} &= \frac{\mu_{xy}}{\sigma_x \cdot \sigma_y} = \frac{A \cdot D + B \cdot E}{\sqrt{(A^2 + B^2)(D^2 + E^2)}}. \end{aligned}$$

Schließen wir die Sonderfälle $A = B = 0$ (d. h. $x \equiv 0$) sowie $D = E = 0$ (d. h. $y \equiv 0$) aus, so liefert die Gleichung stets eindeutige Werte für den Korrelationskoeffizienten im Bereich $-1 \leq \rho_{xy} \leq +1$.

Beispiel: Setzen wir zum Beispiel $A = E = 0$, so ergibt sich der Korrelationskoeffizient $\rho_{xy} = 0$. Dieses Ergebnis ist einsichtig: Nun hängt x nur noch von v und y ausschließlich von u ab. Da aber u und v als statistisch unabhängig angenommen wurden, bestehen keine Beziehungen zwischen x und y . – Ebenso ergibt sich $\rho_{xy} = 0$ für $B = D = 0$.

Die Konstellation $B = E = 0$ führt dazu, dass sowohl x als auch y nur noch von u abhängen. In diesem Fall ergibt sich für den Korrelationskoeffizienten $\rho_{xy} = \pm 1$:

$$\rho_{xy} = \frac{A \cdot D}{\sqrt{A^2 \cdot D^2}} = \frac{A \cdot D}{|A| \cdot |D|} = \pm 1.$$

Besitzen A und D gleiches Vorzeichen, so ist $\rho_{xy} = +1$. Bei unterschiedlichen Vorzeichen ergibt sich der Korrelationskoeffizient -1 . – Auch für $A = D = 0$ ergibt sich der Koeffizient $\rho_{xy} = \pm 1$.

Erzeugung korrelierter Zufallsgrößen

Die Gleichungen der letzten Seite können zur Erzeugung einer zweidimensionalen Zufallsgröße (x, y) mit vorgegebenen Kenngrößen σ_x , σ_y und ρ_{xy} genutzt werden. Wenn außer diesen drei Sollwerten keine weiteren Voraussetzungen getroffen werden, ist einer der vier Koeffizienten A , B , D und E frei wählbar. Im Folgenden wird stets willkürlich $E = 0$ gesetzt.

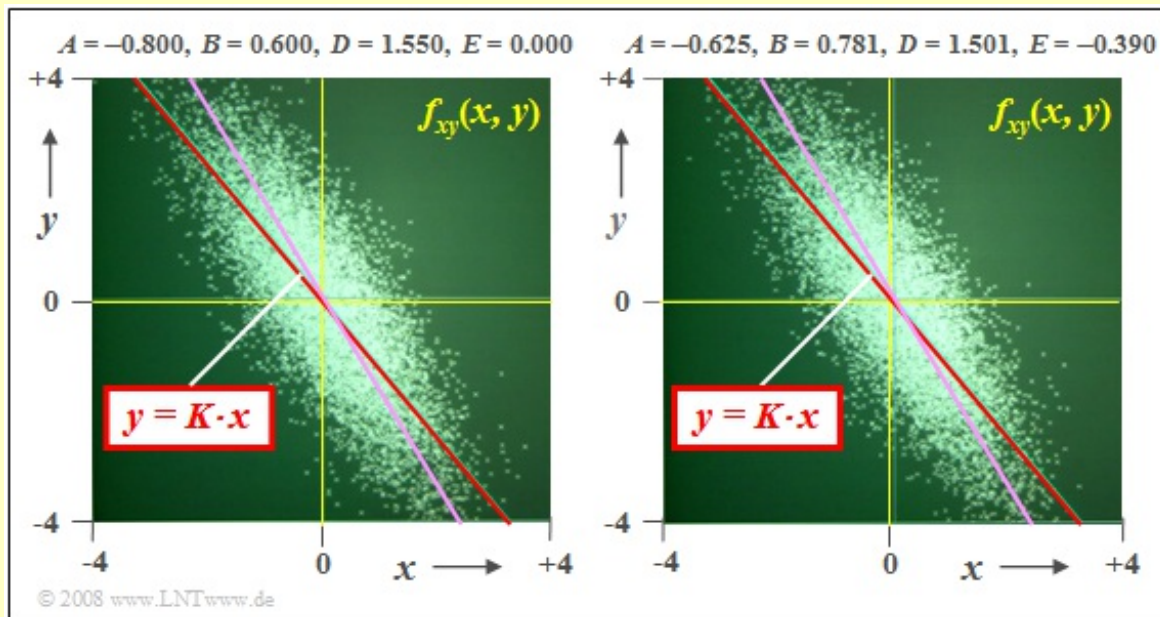
Mit der weiteren Festlegung, dass die statistisch voneinander unabhängigen Zufallsgrößen u und v jeweils die gleiche Streuung $\sigma = 1$ aufweisen, erhält man:

$$D = \sigma_y, \quad A = \sigma_x \cdot \rho_{xy}, \quad B = \sigma_x \cdot \sqrt{1 - \rho_{xy}^2}$$

Bei $\sigma \neq 1$ sind diese Werte jeweils noch durch σ zu dividieren.

Beispiel: Zur Erzeugung einer 2D-Zufallsgröße mit den gewünschten Kennwerten $\sigma_x = 1$, $\sigma_y = 1.55$ und $\rho_{xy} = -0.8$ eignet sich z. B. der Parametersatz $A = -0.8$, $B = 0.6$, $D = 1.55$, $E = 0$, der dem linken Bild zugrundeliegt.

- Die Zufallsgrößen u und v sind dabei gaußförmig und besitzen jeweils die Streuung $\sigma = 1$.
- Die Korrelationsgerade $y = K \cdot x$ (rot dargestellt) verläuft unter einem Winkel von etwa -50 Grad. Violett eingezeichnet ist die Ellipsenhauptachse.



Mit den Parameterwerten $A = -0.625$, $B = 0.781$, $D = 1.501$ und $E = -0.390$ entsprechend der rechten Grafik erhält man – im statistischen Sinne – das gleiche Resultat, auch wenn sich die beiden Punktwolken im Detail unterscheiden.

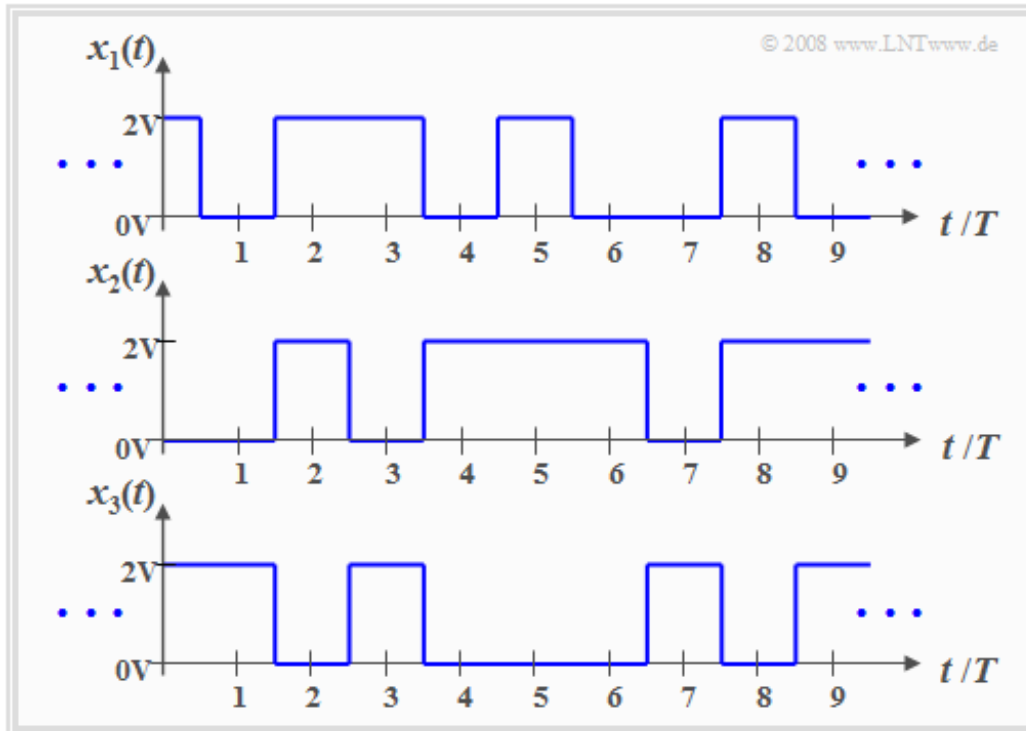
Zufallsprozesse (1)

Ein wichtiger Begriff der stochastischen Signaltheorie ist der **Zufallsprozess**. Nachfolgend sind einige Charakteristika eines solchen *stochastischen Prozesses* – diese Bezeichnungen werden sowohl in der Literatur als auch in unserem Tutorial synonym verwendet – zusammengestellt:

- Unter einem Zufallsprozess $\{x_i(t)\}$ verstehen wir ein *mathematisches Modell* für ein Ensemble von (vielen) Zufallssignalen, die sich zwar im Detail durchaus voneinander unterscheiden können, trotzdem aber gewisse gemeinsame Eigenschaften aufweisen.
- Zur Beschreibung eines Zufallsprozesses $\{x_i(t)\}$ gehen wir von der Vorstellung aus, dass beliebig viele, in ihren physikalischen und statistischen Eigenschaften völlig gleiche Zufallsgeneratoren vorhanden sind, von denen jeder ein Zufallssignal $x_i(t)$ liefert.
- Jeder Zufallsgenerator gibt trotz gleicher physikalischer Realisierung ein anderes Zeitsignal $x_i(t)$ ab, das für alle Zeiten von $-\infty$ bis $+\infty$ existiert. Man bezeichnet dieses spezifische Zufallssignal als das *i-te Mustersignal*.
- Jeder Zufallsprozess beinhaltet mindestens eine stochastische Komponente – zum Beispiel die Amplitude, Frequenz oder Phase eines Nachrichtensignals – und kann daher von einem Beobachter nicht exakt vorausgesagt werden.
- Der Zufallsprozess unterscheidet sich von den sonst in der Wahrscheinlichkeitsrechnung bzw. der Statistik üblichen Zufallsexperimenten dadurch, dass das Ergebnis kein *Ereignis* ist, sondern ein *Funktionsverlauf* (Zeitsignal).
- Betrachtet man den Zufallsprozess $\{x_i(t)\}$ zu einem festen Zeitpunkt, so gelangt man wieder zu dem einfacheren Modell von **Kapitel 2.1**, nach dem das Versuchsergebnis ein Ereignis ist, das einer Zufallsgröße zugeordnet werden kann.

Zufallsprozesse (2)

Die Aussagen der vorherigen Seite werden nun am Beispiel eines binären Zufallsgenerators verdeutlicht, der – zumindest gedanklich – beliebig oft realisiert werden kann.



Die Grafik zeigt drei unterschiedliche Mustersignale mit folgenden Eigenschaften:

- Der hier vorliegende Zufallsprozess $\{x_i(t)\}$ besteht aus einem Ensemble rechteckförmiger Musterfunktionen, die jeweils wie folgt beschrieben werden können:

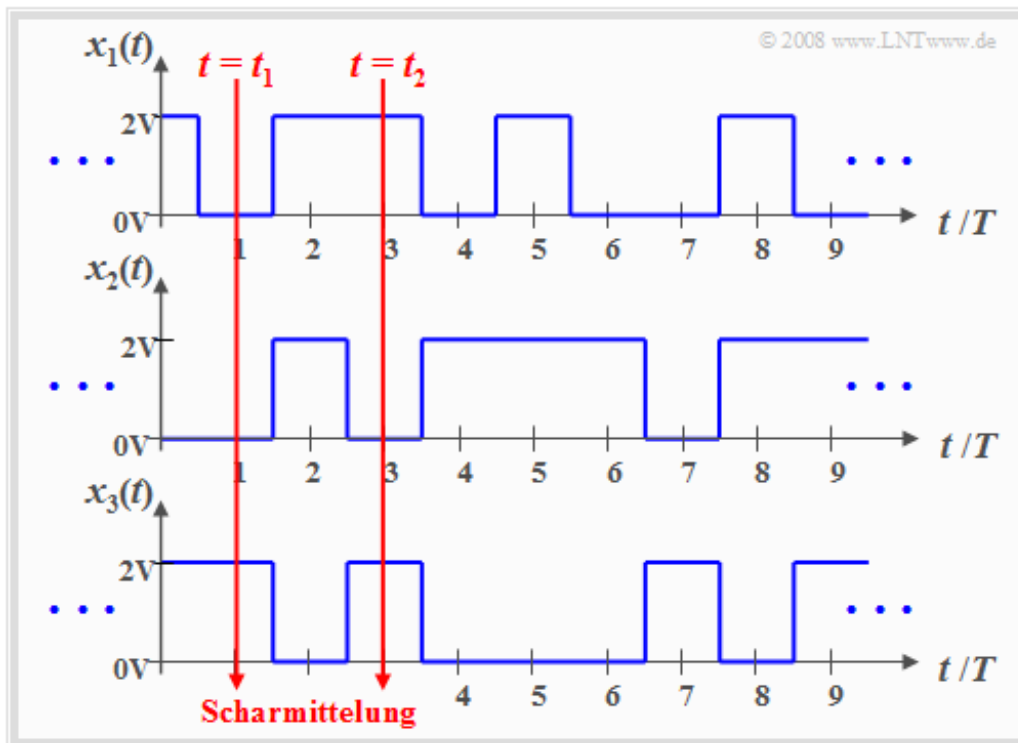
$$x_i(t) = \sum_{\nu=-\infty}^{+\infty} (a_\nu)_i \cdot g(t - \nu \cdot T).$$

- Der Grundimpuls $g(t)$ besitzt im Bereich von $-T/2$ bis $+T/2$ den Wert $2V$; außerhalb ist er 0 . Unter einem *Impuls* verstehen wir gemäß der Definition im Buch „Signaldarstellung“ ein sowohl *deterministisches* als auch *energiebegrenzt*es Signal.
- Die Statistik des hier betrachteten Zufallsprozesses ist somit allein auf die dimensionslosen Amplitudenkoeffizienten $(a_\nu)_i \in \{0, 1\}$ zurückzuführen, die bei der i -ten Musterfunktion mit dem Zeitindex ν versehen sind.
- Trotz der im Detail unterschiedlichen Signalverläufe weisen die skizzierten Mustersignale $x_1(t)$, $x_2(t)$, $x_3(t)$ und auch alle weiteren Mustersignale $x_4(t)$, $x_5(t)$, $x_6(t)$, ... gewisse Gemeinsamkeiten auf, die nachfolgend herausgearbeitet werden sollen.

Stationäre Zufallsprozesse

Definiert man den Momentanwert aller Musterfunktionen $x_i(t)$ zu einem festen Zeitpunkt $t = t_1$ als eine neue Zufallsgröße $x_1 = \{x_i(t_1)\}$, so lassen sich deren statistische Eigenschaften nach den Aussagen von **Kapitel 2** und **Kapitel 3** beschreiben. In gleicher Weise erhalten wir für den Betrachtungszeitpunkt $t = t_2$ die Zufallsgröße $x_2 = \{x_i(t_2)\}$.

Hinweis: Beachten Sie bitte, dass $x_1(t)$ und $x_2(t)$ Musterfunktionen des Zufallsprozesses $\{x_i(t)\}$ sind, während die Zufallsgrößen x_1 und x_2 den Prozess zu den Zeiten t_1 und t_2 charakterisieren.



Die Berechnung der statistischen Kenngrößen muss durch **Scharmittelung** über alle möglichen Musterfunktionen erfolgen (Mittelung über die Laufvariable i , also über alle Realisierungen).

Definition: Bei einem **stationären Zufallsprozess** $\{x_i(t)\}$ sind alle statistischen Kenngrößen der Zufallsgrößen x_1 und x_2 gleich. Auch zu jedem anderen Zeitpunkt ergeben sich genau gleiche Werte.

Die Umkehrung lautet: Man bezeichnet einen Zufallsprozess $\{x_i(t)\}$ als *nichtstationär*, wenn er zu verschiedenen Zeitpunkten unterschiedliche statistische Eigenschaften (Mittelwert, Streuung, Momente höherer Ordnung, Auftrittswahrscheinlichkeiten, etc.) aufweist.

Beispiel: Eine große Anzahl von Mess-Stationen am Äquator ermitteln täglich um 12 Uhr Ortszeit die Temperatur. Mittelt man über all diese Messwerte, so kann man den Einfluss lokaler Indikatoren (z. B. Golfstrom) eliminieren. Trägt man die Mittelwerte (Scharmittelung) über der Zeit auf, so wird sich nahezu eine Konstante ergeben, und man kann von einem *stationären Prozess* sprechen.

Eine vergleichbare Messreihe am 50. Breitengrad würde aufgrund der jahreszeitlichen Schwankungen auf einen *nichtstationären Prozess* hinweisen mit deutlichen Unterschieden hinsichtlich Mittelwert und Varianz der Mittagstemperatur zwischen Januar und Juli.

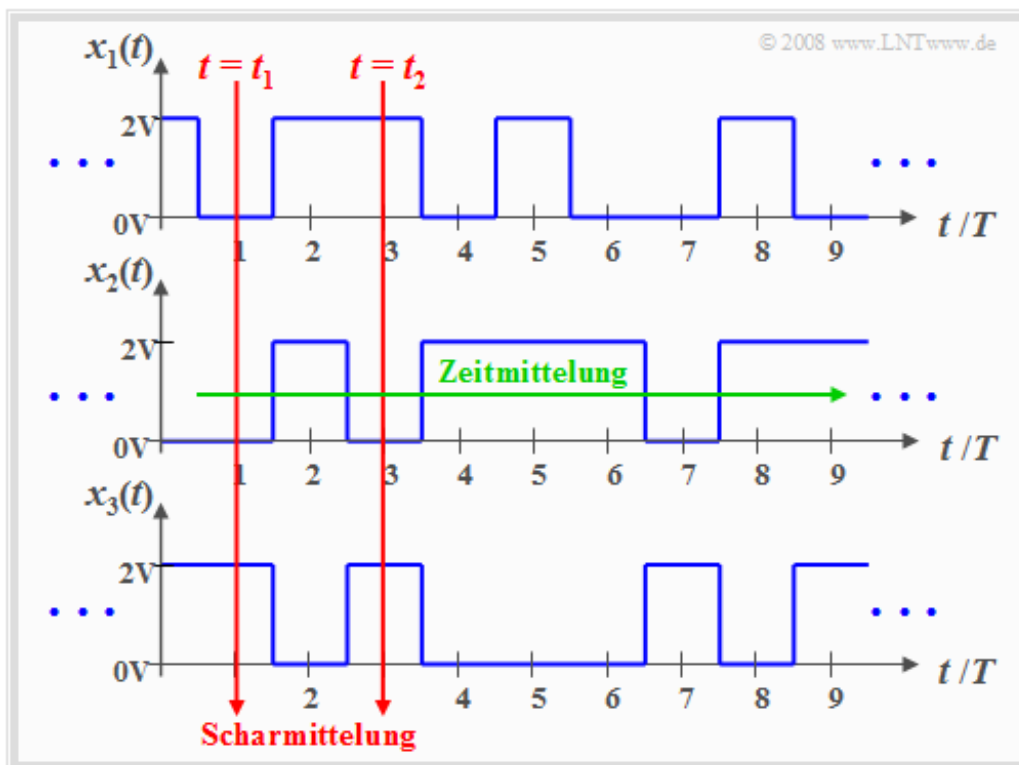
Ergodische Zufallsprozesse

Eine wichtige Unterklasse der stationären Zufallsprozesse sind die so genannten **ergodischen Prozesse** mit folgenden Eigenschaften:

- Bei einem ergodischen Prozess $\{x_i(t)\}$ ist jede einzelne Musterfunktion $x_i(t)$ repräsentativ für das gesamte Ensemble.
- Alle statistischen Beschreibungsgrößen eines ergodischen Prozesses kann man aus einer einzigen Musterfunktion durch Zeitmittelung (bezüglich der Laufvariablen v) gewinnen.
- Das bedeutet auch: Bei Ergodizität stimmen die Zeitmittelwerte einer jeden Musterfunktion mit den entsprechenden Scharmittelwerten zu beliebigen Zeitpunkten überein.
- Beispielsweise gilt bei Ergodizität für das Moment k -ter Ordnung:

$$m_k = \overline{x^k(t)} = E[x^k].$$

Die überstreichende Linie kennzeichnet hierbei den Zeitmittelwert, während der Scharmittelwert durch Erwartungswertwertbildung $E[\dots]$ zu ermitteln ist, wie in **Kapitel 2.2** beschrieben.



Anmerkung: Die Ergodizität lässt sich aus einer endlichen Anzahl von Musterfunktionen und endlichen Signalausschnitten nicht nachweisen. Allerdings wird in den meisten Anwendungen zwar hypothetisch – aber trotzdem durchaus berechtigt – von Ergodizität ausgegangen. Anhand der gefundenen Ergebnisse muss anschließend die Plausibilität dieser **Ergodizitätshypothese** überprüft werden.

Allgemeingültige Beschreibung von Zufallsprozessen

Ist der zu analysierende Zufallsprozess $\{x_i(t)\}$ nicht stationär und damit mit Sicherheit auch nicht ergodisch, so müssen die Momente stets als **Schmitttelwerte** bestimmt werden. Im Allgemeinen sind diese zeitabhängig:

$$m_k(t_1) \neq m_k(t_2).$$

Da durch die Momente aber auch die **charakteristische Funktion** (Fourierrücktransformierte der WDF)

$$C_x(\Omega) = \sum_{k=0}^{\infty} \frac{m_k}{k!} \cdot \Omega^k \quad \circ \longrightarrow \bullet \quad f_x(x)$$

festliegt, ist die Wahrscheinlichkeitsdichtefunktion $f_x(x)$ ebenfalls zeitabhängig.

Sollen nun nicht nur die Amplitudenverteilungen zu den verschiedenen Zeitpunkten t_1, t_2, \dots ermittelt werden, sondern auch die statistischen Bindungen zwischen den Signalwerten zu diesen Zeiten, so muss man auf die **zweidimensionale Verbundwahrscheinlichkeitsdichtefunktion** übergehen.

Betrachtet man beispielsweise die beiden Zeitpunkte t_1 und t_2 , so ist folgendes zu vermerken:

- Die 2D-WDF ergibt sich entsprechend den Angaben in **Kapitel 4.1** mit $x = x(t_1)$ und $y = x(t_2)$. Es ist offensichtlich, dass bereits die Ermittlung dieser Größe sehr aufwändig ist.
- Berücksichtigt man weiter, dass zur Erfassung aller statistischen Bindungen innerhalb des Zufallsprozesses eigentlich die n -dimensionale Verbundwahrscheinlichkeitsdichtefunktion (VWDF) herangezogen werden müsste, wobei möglichst noch der Grenzwert $n \rightarrow \infty$ zu bilden ist, so erkennt man die Schwierigkeiten für die Lösung praktischer Probleme.
- Aus diesen Gründen geht man zur Beschreibung der statistischen Bindungen eines Zufallsprozesses auf die Autokorrelationsfunktion über, die nachfolgend definiert wird.

Allgemeine Definition der Autokorrelationsfunktion

Definition: Die **Autokorrelationsfunktion** (AKF) eines beliebigen Zufallsprozesses $\{x_i(t)\}$ ist gleich dem Erwartungswert des Produkts der Signalwerte zu zwei Zeitpunkten t_1 und t_2 :

$$\varphi_x(t_1, t_2) = E[x(t_1) \cdot x(t_2)].$$

Diese Definition gilt unabhängig davon, ob der Zufallsprozess ergodisch oder nichtergodisch ist, und sie gilt auch für nichtstationäre Prozesse.

Ein Vergleich mit **Kapitel 4.1** zeigt, dass der AKF-Wert $\varphi_x(t_1, t_2)$ das gemeinsame Moment m_{11} zwischen den beiden Zufallsgrößen $x(t_1)$ und $x(t_2)$ angibt.

Hinweis: Um den Zusammenhang mit der **Kreuzkorrelationsfunktion** φ_{xy} zwischen den beiden statistischen Größen x und y deutlich zu machen (siehe Kapitel 4.6), wird in mancher Literatur für die AKF auch häufig die Schreibweise $\varphi_{xx}(t_1, t_2)$ verwendet.

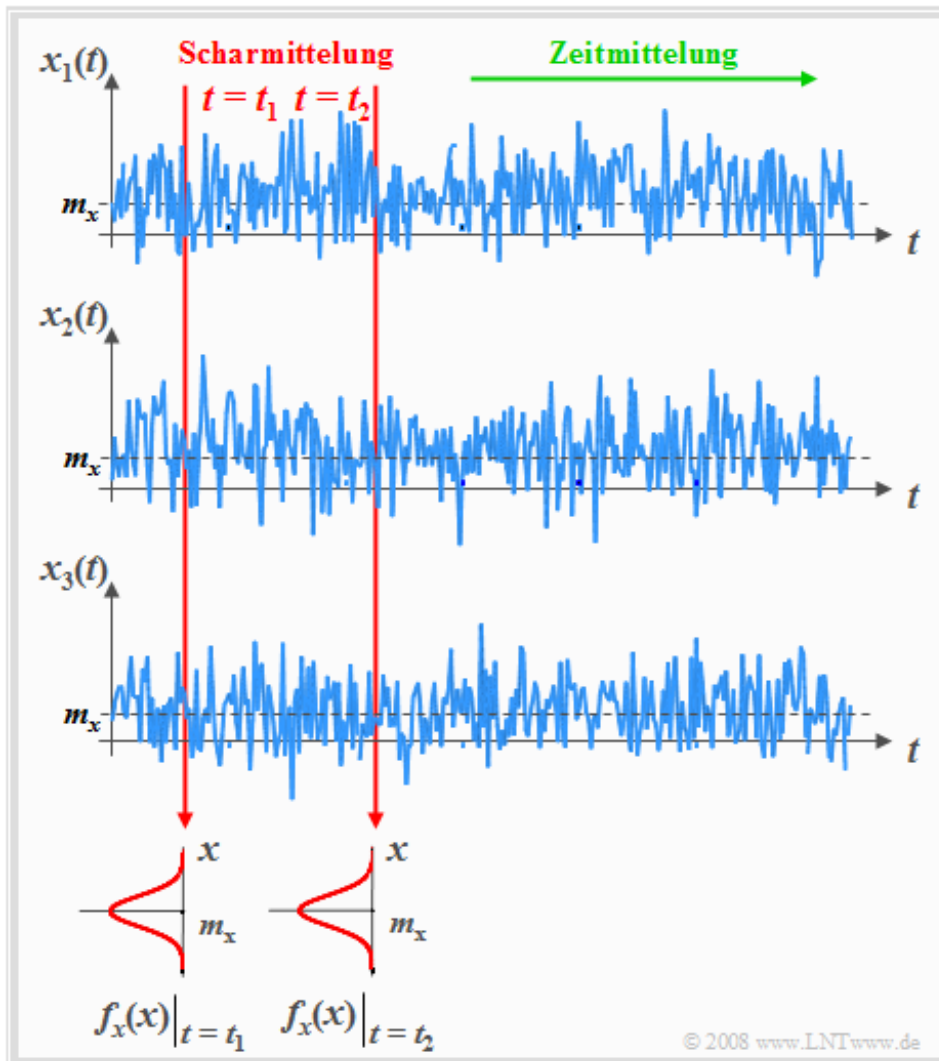
Während für exakte Aussagen hinsichtlich der statistischen Bindungen eines Zufallsprozesses eigentlich die n -dimensionale Verbunddichte (mit $n \rightarrow \infty$) benötigt wird, werden durch den Übergang auf die Autokorrelationsfunktion implizit folgende Vereinfachungen getroffen:

- Anstelle von unendlich vielen Zeitpunkten werden hier nur zwei betrachtet.
- Anstelle aller Momente m_{kl} zu den beiden Zeitpunkten t_1 und t_2 mit $k, l \in \{1, 2, 3, \dots\}$ wird hier nur das gemeinsame Moment m_{11} erfasst.
- Das Moment m_{11} gibt ausschließlich die lineare Abhängigkeit (Korrelation) des Prozesses wieder. Alle statistische Bindungen höherer Ordnung werden dagegen nicht berücksichtigt.
- Deshalb sollte bei der Bewertung von Zufallsprozessen mittels AKF stets berücksichtigt werden, dass diese nur sehr beschränkte Aussagen über die statistischen Bindungen erlaubt.

Die obige Definition der Autokorrelationsfunktion gilt allgemein, also auch für nichtstationäre und nichtergodische Prozesse. Ein Beispiel eines nichtstationären Vorgangs ist das Auftreten von Impulsstörungen im Fernsprechnet, verursacht durch Wählimpulse in benachbarten Leitungen. Bei Digitalübertragung führen solche nichtstationären Störprozesse meist zu Bündelfehlern.

Autokorrelationsfunktion bei ergodischen Prozessen (1)

Im Folgenden beschränken wir uns auf stationäre und ergodische Prozesse. Ein solcher Prozess $\{x_i(t)\}$ wird zum Beispiel bei der Untersuchung von **Thermischem Rauschen** zugrunde gelegt. Dabei wird von der Vorstellung ausgegangen, dass beliebig viele, in ihren physikalischen und statistischen Eigenschaften völlig gleiche Widerstände vorhanden sind, von denen jeder ein anderes Zufallssignal $x_i(t)$ abgibt.



Zu dieser Grafik ist anzumerken:

- Die einzelnen Musterfunktionen $x_i(t)$ können zu allen beliebigen Zeiten alle beliebigen Werte annehmen. Das bedeutet, dass der hier betrachtete Zufallsprozess $\{x_i(t)\}$ sowohl wert- als auch zeitkontinuierlich ist.
- Auch wenn über die tatsächlichen Signalwerte der einzelnen Musterfunktionen aufgrund der Stochastik keine Aussagen getroffen werden können, so sind die Momente und die WDF zu allen Zeitpunkten gleich.
- In obiger Grafik ist aus Gründen einer verallgemeinerten Darstellung auch ein Gleichanteil m_x berücksichtigt, der bei Thermischem Rauschen nicht vorhanden ist.

Autokorrelationsfunktion bei ergodischen Prozessen (2)

Definition: Von einem **stationären Zufallsprozess** $\{x_i(t)\}$ spricht man, wenn seine statistischen Eigenschaften invariant gegenüber Zeitverschiebungen sind. Für die Autokorrelationsfunktion (AKF) bedeutet diese Aussage, dass sie nicht mehr eine Funktion der beiden unabhängigen Zeitvariablen t_1 und t_2 ist, sondern nur noch von der Zeitdifferenz $\tau = t_2 - t_1$ abhängt:

$$\varphi_x(t_1, t_2) \rightarrow \varphi_x(\tau) = E[x(t) \cdot x(t + \tau)].$$

Die Scharmittelung kann dabei zu jeder beliebigen Zeit t erfolgen.

Unter der weiteren Annahme eines ergodischen Zufallsprozesses können alle Momente auch durch Zeitmittelung über eine einzige ausgewählte Musterfunktion $x(t)$ ermittelt werden. Diese Zeitmittelwerte stimmen mit den entsprechenden Scharmittelwerten überein.

Damit folgt für die **AKF eines ergodischen Prozesses**, dessen Mustersignale jeweils von $-\infty$ bis $+\infty$ reichen (T_M bezeichnet die Messdauer):

$$\varphi_x(\tau) = \overline{x(t) \cdot x(t + \tau)} = \lim_{T_M \rightarrow \infty} \frac{1}{T_M} \cdot \int_{-T_M/2}^{T_M/2} x(t) \cdot x(t + \tau) dt.$$

Die überstreichende Linie kennzeichnet die Zeitmittelung über das unendlich ausgedehnte Zeitintervall.

Bei periodischen Signalen kann man auf den Grenzübergang verzichten, so dass in diesem Sonderfall die Autokorrelationsfunktion mit der Periodendauer T_0 auch in folgender Weise geschrieben werden kann:

$$\varphi_x(\tau) = \frac{1}{T_0} \cdot \int_{-T_0/2}^{T_0/2} x(t) \cdot x(t + \tau) dt = \frac{1}{T_0} \cdot \int_0^{T_0} x(t) \cdot x(t + \tau) dt.$$

Wichtig ist dabei nur, dass insgesamt genau über eine Periodendauer T_0 (oder Vielfache davon) gemittelt wird. Es ist egal, welcher Zeitausschnitt dabei verwendet wird.

Eigenschaften der Autokorrelationsfunktion

Nachfolgend sind wichtige Eigenschaften der Autokorrelationsfunktion (AKF) zusammengestellt:

- Ist der betrachtete Zufallsprozess reell, so gilt dies auch für seine AKF.
- Die AKF besitzt die Einheit einer Leistung, beispielsweise Watt („W“). Häufig bezieht man diese auf den Einheitswiderstand 1Ω ; in diesem Fall hat $\varphi_x(\tau)$ die Einheit „V²“ bzw. „A²“.
- Die AKF ist immer eine gerade Funktion $\Rightarrow \varphi_x(-\tau) = \varphi_x(\tau)$. Alle Phasenbeziehungen des Zufallsprozesses gehen in der AKF verloren.
- Die AKF an der Stelle $\tau = 0$ gibt den quadratischen Mittelwert m_2 (Moment 2. Ordnung) und damit die gesamte Signalleistung (Gleich- und Wechselanteil) an:

$$\varphi_x(0) = m_2 = \overline{x^2(t)}.$$

- Das AKF-Maximum ist stets bei $\tau = 0$, und es gilt: $|\varphi_x(\tau)| \leq \varphi_x(0)$. Bei nichtperiodischen Prozessen ist für $\tau \neq 0$ der Betrag $|\varphi_x(\tau)|$ stets kleiner als die Leistung $\varphi_x(0)$.
- Bei einem periodischen Zufallsprozess weist die AKF die gleiche Periodendauer T_0 wie die einzelnen Mustersignale $x_i(t)$ auf:

$$\varphi_x(\pm T_0) = \varphi_x(\pm 2 \cdot T_0) = \dots = \varphi_x(0).$$

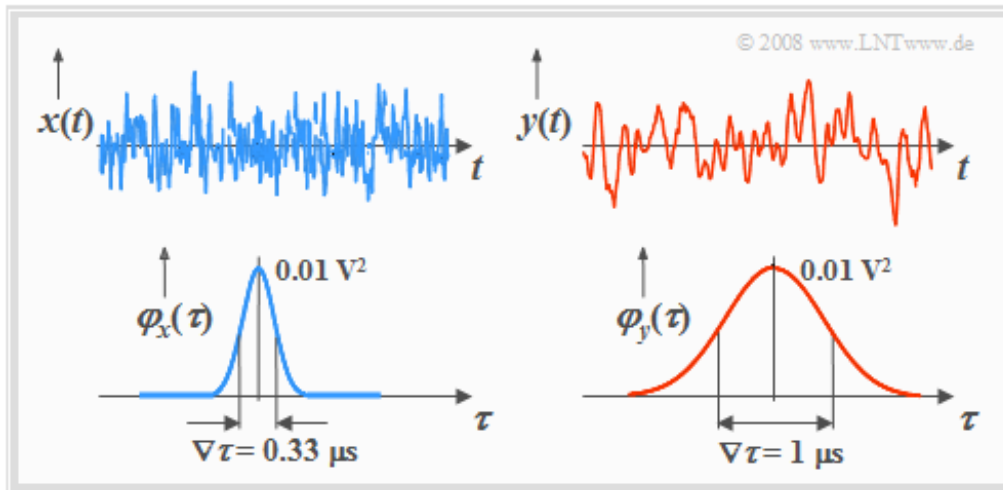
- Der Gleichanteil m_1 eines nichtperiodischen Signals kann aus dem Grenzwert der AKF für $\tau \rightarrow \infty$ berechnet werden. Hierbei gilt:

$$\lim_{\tau \rightarrow \infty} \varphi_x(\tau) = m_1^2 = \overline{[x(t)]}^2.$$

- Dagegen schwankt bei Signalen mit periodischen Anteilen der Grenzwert der AKF für $\tau \rightarrow \infty$ um diesen Endwert (Quadrat des Gleichanteils).

Interpretation der Autokorrelationsfunktion

Das nachfolgende Bild zeigt oben je ein Mustersignal zweier verschiedener Prozesse $\{x_i(t)\}$ und $\{y_i(t)\}$, unten die zugehörigen Autokorrelationsfunktionen. Der Prozess $\{y_i(t)\}$ weist stärkere innere statistische Bindungen auf als $\{x_i(t)\}$. Spektral gesehen ist er niederfrequenter.



Anhand dieser Grafiken sind folgende Aussagen möglich:

- Die beiden Mustersignale lassen bereits vermuten, dass beide Prozesse mittelwertfrei sind den gleichen Effektivwert aufweisen.
- Anhand der Autokorrelationsfunktionen werden diese Aussagen bestätigt. Die leeren Mittelwerte $m_x = m_y = 0$ ergeben sich jeweils aus dem AKF-Grenzwert für $\tau \rightarrow \infty$.
- Wegen $m_x = 0$ gilt für die Varianz: $\sigma_x^2 = \varphi_x(0) = 0.01 \text{ V}^2$, und der Effektivwert ist somit $\sigma_x = 0.1 \text{ V}$. Das Signal $y(t)$ weist die gleiche Varianz und den gleichen Effektivwert auf.
- Die AKF-Werte fallen um so langsamer ab, je stärker die inneren statistischen Bindungen sind. Während sich das Signal $x(t)$ mit relativ schmaler AKF zeitlich sehr schnell ändert, reichen bei dem niederfrequenteren Signal $y(t)$ die statistischen Bindungen deutlich weiter.
- Das bedeutet aber auch, dass der Signalwert $y(t + \tau)$ aus $y(t)$ besser vorhergesagt werden kann als $x(t + \tau)$ aus $x(t)$.
- Eine quantitative Kenngröße für die Stärke der statistischen Bindungen ist die **äquivalente AKF-Dauer** $\nabla\tau$, die sich aus der AKF über das flächengleiche Rechteck ermitteln lässt:

$$\nabla\tau = \frac{1}{\varphi_x(0)} \cdot \int_{-\infty}^{\infty} \varphi_x(\tau) \, d\tau. \quad \text{Hinweis: Man spricht „Nabla-tau“}.$$

Beim hier betrachteten Prozess (mit gaußähnlicher AKF) ist $\nabla\tau = 0.33 \mu\text{s}$ bzw. $\nabla\tau = 1 \mu\text{s}$.

- Als ein weiteres Maß für die Stärke der statistischen Bindungen wird in der Literatur häufig die **Korrelationsdauer** T_K verwendet. Diese gibt die Zeitdauer an, bei der die AKF auf die Hälfte ihres Maximalwertes abgefallen ist.

Numerische AKF-Ermittlung (1)

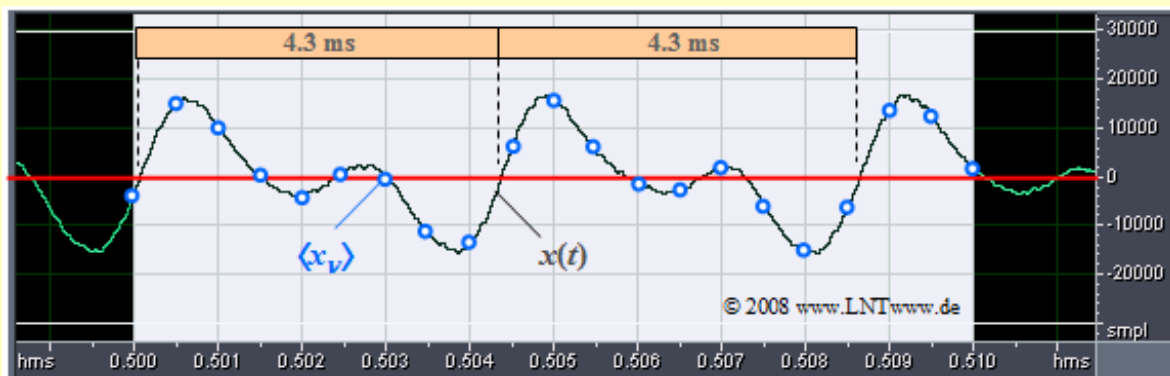
Bisher haben wir stets zeitkontinuierliche Signale $x(t)$ betrachtet, die für die Darstellung und Simulation mittels Digitalrechner ungeeignet sind. Hierzu ist eine **zeitdiskrete Signaldarstellung** $\langle x_\nu \rangle$ erforderlich, wie im **Kapitel 5.1** von Buch „Signaldarstellung“ dargelegt. Hier eine kurze Zusammenfassung:

Das zeitdiskrete Signal $\langle x_\nu \rangle$ ist die Folge der Abtastwerte $x_\nu = x(\nu \cdot T_A)$. Das zeitkontinuierliche Signal $x(t)$ wird durch die Folge $\langle x_\nu \rangle$ vollständig beschrieben, wenn das **Abtasttheorem** erfüllt ist:

$$T_A \leq \frac{1}{2 \cdot B_x}$$

B_x bezeichnet hierbei die absolute (einseitige) Bandbreite des Analogsignals $x(t)$. Diese sagt aus, dass die Spektralfunktion $X(f)$ für alle Frequenzen $|f| > B_x$ gleich 0 ist.

Beispiel: Das Bild zeigt einen Ausschnitt eines Audiosignals der Dauer 10 Millisekunden. Obwohl das gesamte Signal ein breites Spektrum mit der Mittenfrequenz bei etwa 500 Hz besitzt, ist während des betrachteten Zeitintervalls ein (nahezu) periodisches Signal mit Periodendauer $T_0 = 4.3$ Millisekunden zu erkennen. Daraus ergibt sich die Grundfrequenz zu etwa $f_0 = 230$ Hz.



Blau eingezeichnet sind die Abtastwerte im Abstand $T_A = 0.5$ ms. Diese Folge $\langle x_\nu \rangle$ von Abtastwerten würde die gesamte Information über das Signal $x(t)$ beinhalten, wenn das Analogsignal $x(t)$ auf den Frequenzbereich bis 1 kHz begrenzt wäre. Sind im Signal $x(t)$ höhere Frequenzanteile enthalten, so muss T_A entsprechend kleiner gewählt werden.

Da nun die Signalwerte nur zu diskreten Zeitpunkten (bei Vielfachen von T_A) vorliegen, kann man auch die Autokorrelationsfunktion nur zu ganzzahligen Vielfachen von T_A bestimmen. Mit den zeitdiskreten Signalwerten $x_\nu = x(\nu \cdot T_A)$ und $x_{\nu+k} = x((\nu+k) \cdot T_A)$ sowie der zeitdiskreten AKF $\varphi_k = \varphi_x(k \cdot T_A)$ lässt sich somit die AKF-Berechnung wie folgt darstellen:

$$\varphi_k = \overline{x_\nu \cdot x_{\nu+k}}$$

Die überstreichende Linie kennzeichnet hierbei wieder die Zeitmittelung.

Numerische AKF-Ermittlung (2)

Wir stellen uns nun die Aufgabe, die AKF-Stützstellen $\varphi_0, \dots, \varphi_l$ aus N Abtastwerten (x_1, \dots, x_N) zu ermitteln, wobei der Parameter l sehr viel kleiner als N vorausgesetzt wird. Beispielsweise gelte $l = 100$ und $N = 100000$.

Die AKF-Berechnungsvorschrift lautet nun (mit $0 \leq k \leq l$):

$$\varphi_k = \frac{1}{N - k} \cdot \sum_{\nu=1}^{N-k} x_\nu \cdot x_{\nu+k}.$$

Bringen wir $N - k$ auf die linke Seite, so erhalten wir daraus $l + 1$ Gleichungen, nämlich:

$$k = 0 : \quad N \cdot \varphi_0 \quad = \quad x_1 \cdot x_1 \quad + \quad x_2 \cdot x_2 \quad + \dots \quad + \quad x_\nu \cdot x_\nu \quad + \dots \quad + \quad x_N \cdot x_N,$$

$$k = 1 : \quad (N - 1) \cdot \varphi_1 \quad = \quad x_1 \cdot x_2 \quad + \quad x_2 \cdot x_3 \quad + \dots \quad + \quad x_\nu \cdot x_{\nu+1} \quad + \dots \quad + \quad x_{N-1} \cdot x_N,$$

.....

$$k \text{ allg. : } (N - k) \cdot \varphi_k \quad = \quad x_1 \cdot x_{1+k} \quad + \quad x_2 \cdot x_{2+k} \quad + \dots \quad + \quad x_\nu \cdot x_{\nu+k} \quad + \dots \quad + \quad x_{N-k} \cdot x_N,$$

.....

$$k = l : \quad (N - l) \cdot \varphi_l \quad = \quad x_1 \cdot x_{1+l} \quad + \quad x_2 \cdot x_{2+l} \quad + \dots \quad + \quad x_\nu \cdot x_{\nu+l} \quad + \dots \quad + \quad x_{N-l} \cdot x_N.$$

Hieraus ergibt sich der folgende Algorithmus:

- Man definiert das Feld $\text{AKF}[0 : l]$ vom Typ *float* und belegt alle Elemente mit Nullen vor.
- Bei jedem Schleifendurchlauf (indiziert mit der Variablen k) werden die $l + 1$ Feldelemente $\text{AKF}[k]$ jeweils um den Betrag $x_\nu \cdot x_{\nu+k}$ erhöht.
- Alle $l + 1$ Feldelemente werden allerdings nur dann bearbeitet, so lange die Laufvariable k nicht größer als $N - l$ ist. Es ist stets zu berücksichtigen, dass $\nu + k \leq N$ gelten muss. Das bedeutet, dass die Mittelung in den unterschiedlichen Feldern $\text{AKF}[0] \dots \text{AKF}[l]$ über eine unterschiedliche Anzahl von Summanden erfolgt.
- Werden am Ende der Berechnung noch die in $\text{AKF}[k]$ gespeicherten Werte durch die Anzahl der Summanden $(N - k)$ dividiert, so enthält dieses Feld die gesuchten diskreten AKF-Werte:

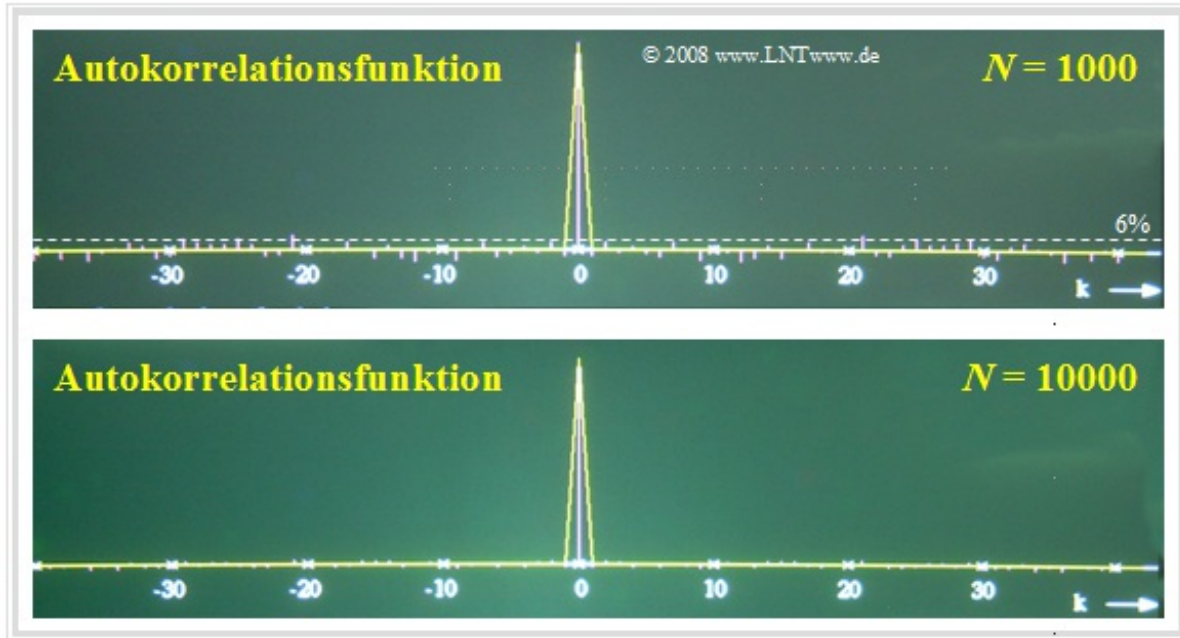
$$\varphi_x(k \cdot T_A) = \text{AKF}[k].$$

Anmerkung: Bei $l \ll N$ kann man den Algorithmus vereinfachen, indem die Anzahl der Summanden für alle k -Werte gleich gewählt werden:

$$\varphi_k = \frac{1}{N - l} \cdot \sum_{\nu=1}^{N-l} x_\nu \cdot x_{\nu+k}.$$

Genauigkeit der numerischen AKF-Berechnung (1)

Der entscheidende Parameter für die Qualität der numerischen AKF-Berechnung ist die Anzahl N der berücksichtigten Abtastwerte. Im nachfolgenden oberen Bild sehen Sie das Ergebnis für $N = 1000$ und darunter für $N = 10000$ Zufallsgrößen.

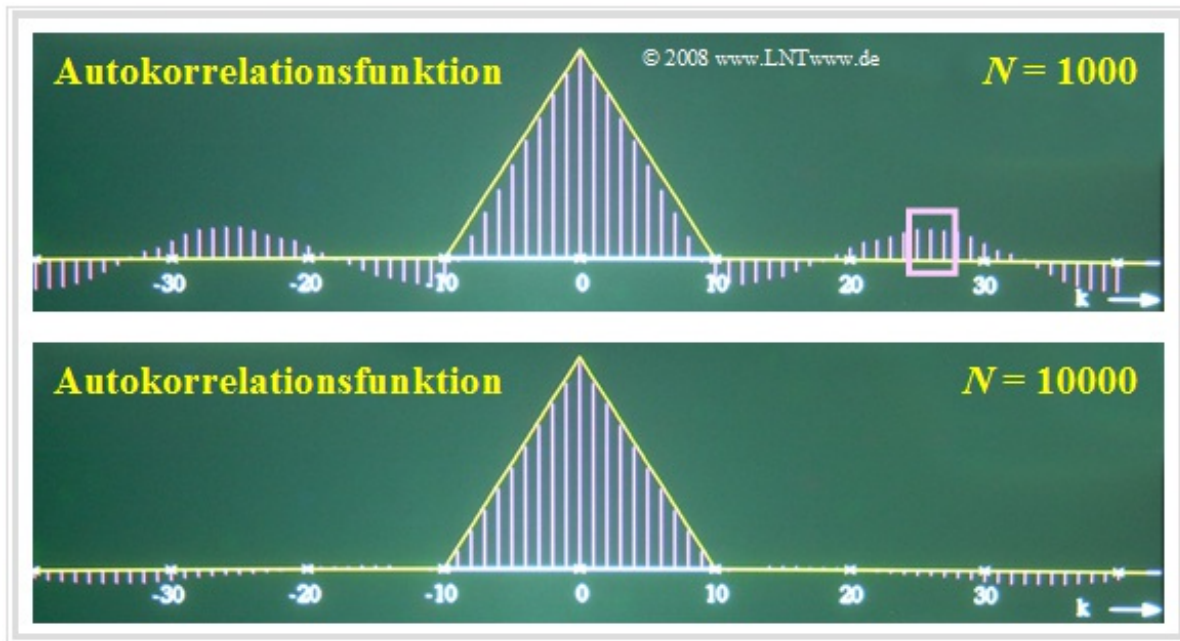


Die betrachteten Zufallsgrößen sind hier voneinander statistisch unabhängig. Somit sollten eigentlich alle AKF-Werte mit Ausnahme des Wertes bei $k = 0$ identisch Null sein. Bei $N = 10000$ (untere Grafik) beträgt der maximale Fehler nur etwa 1% und ist bei dieser Darstellung fast gar nicht sichtbar. Dagegen wächst der Fehler bei $N = 1000$ bis auf $\pm 6\%$ an (Markierung beachten).

Genauigkeit der numerischen AKF-Berechnung (2)

Die Ergebnisse ändern sich, wenn eine Zufallsgröße mit inneren statistischen Bindungen vorliegt. Betrachten wir beispielsweise eine dreieckförmige AKF mit $\varphi_x(k) \neq 0$ für $|k| \leq 10$, so erkennt man deutlich größere Abweichungen, nämlich Fehler bis zu

- etwa $\pm 15\%$ bei $N = 1000$,
- etwa $\pm 5\%$ bei $N = 10000$.



Begründung des schlechteren Ergebnisses:

- Aufgrund der inneren statistischen Bindungen liefern nun nicht mehr alle Abtastwerte die volle Information über den zugrundeliegenden Zufallsprozess.
- Außerdem lassen die Bilder erkennen, dass bei der numerischen AKF-Berechnung einer Zufallsgröße mit statistischen Bindungen auch die Fehler korreliert sind.
- Ist – wie beispielsweise im oberen Bild zu sehen – der AKF-Wert $\varphi_x(26)$ fälschlicherweise positiv und groß, so ergeben sich auch die benachbarten AKF-Werte $\varphi_x(25)$ und $\varphi_x(27)$ als positiv und mit ähnlichen Zahlenwerten. Dieser Bereich ist durch das Rechteck markiert.

Theorem von Wiener-Chintchine

Im Weiteren beschränken wir uns auf ergodische Prozesse. Wie im **Kapitel 4.4** gezeigt wurde, gelten dann die folgenden Aussagen:

- Jede einzelne Musterfunktion $x_i(t)$ ist repräsentativ für den gesamten Zufallsprozess $\{x_i(t)\}$. Alle Zeitmittelwerte sind somit identisch mit den dazugehörigen Scharmittelwerten.
- Die Autokorrelationsfunktion, die allgemein von den beiden Zeitparametern t_1 und t_2 beeinflusst wird, hängt nur noch von der Zeitdifferenz $\tau = t_2 - t_1$ ab:

$$\varphi_x(t_1, t_2) = E[x(t_1) \cdot x(t_2)] = \varphi_x(\tau) = \int_{-\infty}^{+\infty} x(t) \cdot x(t + \tau) dt.$$

Diese Funktion liefert quantitative Aussagen über die (linearen) statistischen Bindungen innerhalb des ergodischen Prozesses $\{x_i(t)\}$ im Zeitbereich. Die äquivalente Beschreibungsgröße im Frequenzbereich ist die *spektrale Leistungsdichte*, häufig auch als *Leistungsdichtespektrum* (LDS) bezeichnet.

Definition: Das **Leistungsdichtespektrum** (LDS) eines ergodischen Zufallsprozesses $\{x_i(t)\}$ ist die Fouriertransformierte der Autokorrelationsfunktion (AKF):

$$\Phi_x(f) = \int_{-\infty}^{+\infty} \varphi_x(\tau) \cdot e^{-j\pi f\tau} d\tau.$$

Diesen Funktionalzusammenhang nennt man das Theorem von **Wiener** und **Chintchine**.

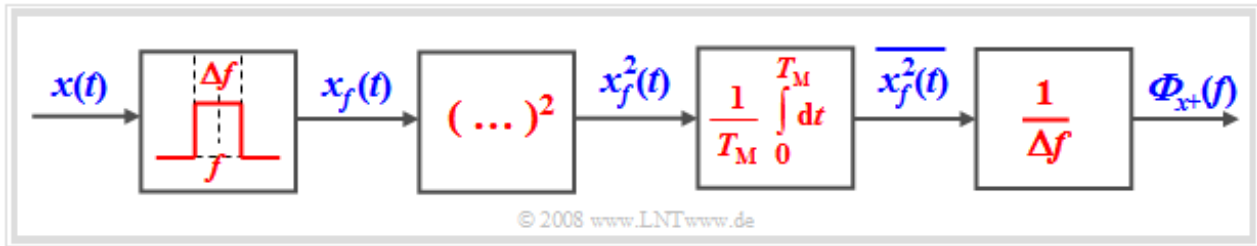
Ebenso kann die AKF als Fourierreücktransformierte des LDS berechnet werden (siehe **Kapitel 3.1** des Buches „Signaldarstellung“):

$$\varphi_x(\tau) = \int_{-\infty}^{+\infty} \Phi_x(f) \cdot e^{j\pi f\tau} df.$$

Die beiden Gleichungen sind nur dann direkt anwendbar, wenn der Zufallsprozess weder einen Gleichanteil noch periodische Anteile beinhaltet. Andernfalls muss man nach den Angaben auf Seite 4 dieses Abschnitts vorgehen: **Spektrale Leistungsdichte mit Gleichsignalkomponente**.

Physikalische Interpretation und Messung

Das folgende Bild zeigt eine Anordnung zur (näherungsweise) messtechnischen Bestimmung des Leistungsdichtespektrums $\Phi_x(f)$.



Hierzu ist folgendes anzumerken:

- Das Zufallssignal $x(t)$ wird auf ein (möglichst) rechteckförmiges und (möglichst) schmalbandiges Filter mit Mittenfrequenz f und Bandbreite Δf gegeben, wobei Δf entsprechend der gewünschten Frequenzauflösung hinreichend klein gewählt werden muss.
- Das entsprechende Ausgangssignal $x_f(t)$ wird quadriert und anschließend der Mittelwert über eine hinreichend lange Messdauer T_M gebildet. Damit erhält man die Leistung von $x_f(t)$ bzw. die Leistungsanteile von $x(t)$ im Spektralbereich von $f - \Delta f/2$ bis $f + \Delta f/2$:

$$P_{xf} = \overline{x_f(t)^2} = \frac{1}{T_M} \cdot \int_0^{T_M} x_f(t)^2 dt.$$

- Die Division durch Δf führt von der spektralen Leistung zur spektralen Leistungsdichte:

$$\Phi_{x+}(f) = \frac{P_{xf}}{\Delta f} \quad \text{bzw.} \quad \Phi_x(f) = \frac{P_{xf}}{2 \cdot \Delta f}.$$

Hierbei bezeichnet $\Phi_{x+}(f) = 2 \cdot \Phi_x(f)$ das einseitige, nur für positive Frequenzen definierte LDS. Für negative Frequenzen ist $\Phi_{x+}(f) = 0$. Im Gegensatz dazu gilt für das üblicherweise verwendete zweiseitige LDS: $\Phi_x(-f) = \Phi_x(f)$.

- Während die Leistung P_{xf} mit kleiner werdender Bandbreite Δf gegen Null tendiert, bleibt die spektrale Leistungsdichte ab einem hinreichend kleinen Wert von Δf nahezu konstant.
- Für die exakte Bestimmung von $\Phi_x(f)$ sind zwei Grenzübergänge notwendig:

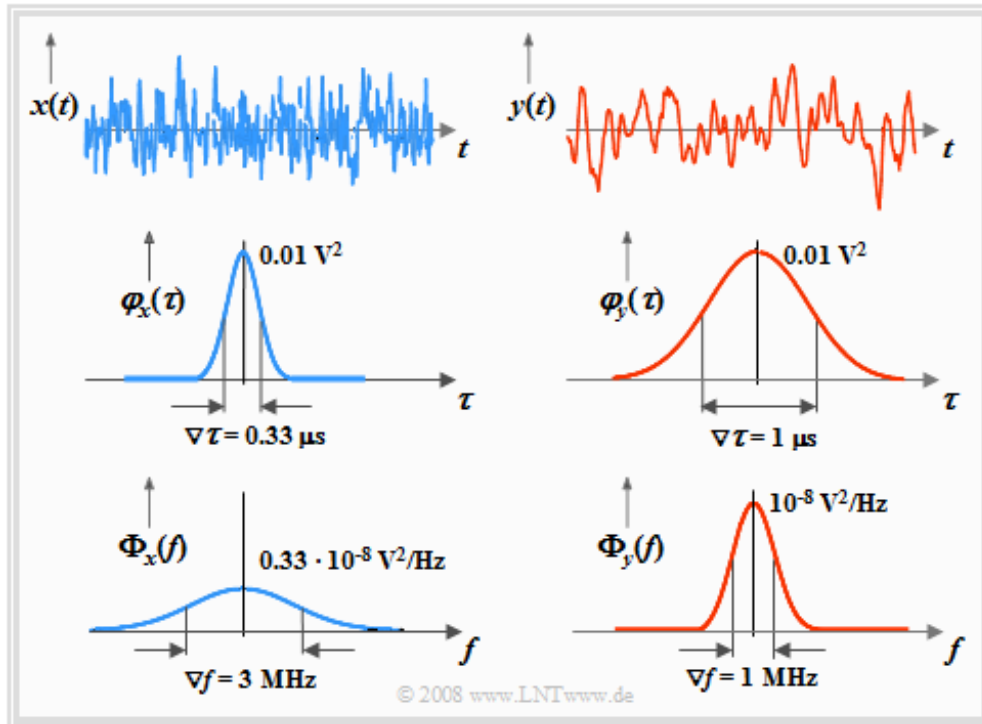
$$\Phi_x(f) = \lim_{\Delta f \rightarrow 0} \lim_{T_M \rightarrow \infty} \frac{1}{2 \cdot \Delta f \cdot T_M} \cdot \int_0^{T_M} x_f^2(t) dt.$$

Aus dieser physikalischen Interpretation folgt weiter, dass das LDS stets reell ist und nie negativ werden kann. Die gesamte Signalleistung von $x(t)$ erhält man dann durch Integration über alle Spektralanteile:

$$P_x = \int_0^{\infty} \Phi_{x+}(f) df = \int_{-\infty}^{+\infty} \Phi_x(f) df.$$

Reziprozitätsgesetz von AKF-Zeitdauer und LDS-Bandbreite (1)

Alle **Gesetzmäßigkeiten der Fouriertransformation**, hergeleitet im Kapitel 3.2 des Buches „Signaldarstellung“ für deterministische Signale, können auch auf die Autokorrelationsfunktion (AKF) und das Leistungsdichtespektrum (LDS) eines Zufallsprozesses angewendet werden. Aufgrund der spezifischen Eigenschaften von AKF (stets reell und gerade) und LDS (stets reell, gerade und nicht-negativ) liefern allerdings nicht alle Gesetze sinnvolle Ergebnisse.



Wir betrachten nun wie auf der Seite **Interpretation der Autokorrelationsfunktion** im Kapitel 4.4 zwei unterschiedliche ergodische Zufallsprozesse $\{x_i(t)\}$ und $\{y_i(t)\}$ anhand

- der beiden Mustersignale $x(t)$ bzw. $y(t) \Rightarrow$ obere Skizze,
- der beiden Autokorrelationsfunktionen $\varphi_x(\tau)$ bzw. $\varphi_y(\tau) \Rightarrow$ mittlere Skizze,
- der beiden Leistungsdichtespektren $\Phi_x(f)$ bzw. $\Phi_y(f) \Rightarrow$ untere Skizze.

Die Interpretation dieser Grafiken folgt auf der nächsten Seite.

Reziprozitätsgesetz von AKF-Zeitdauer und LDS-Bandbreite (2)

Anhand der **Grafiken** auf der letzten Seite sind folgende Aussagen möglich:

- Die Flächen unter den LDS-Kurven sind gleich \Rightarrow die Prozesse besitzen gleiche Leistung:

$$\varphi_x(0) = \int_{-\infty}^{+\infty} \Phi_x(f) df = \varphi_y(0) = \int_{-\infty}^{+\infty} \Phi_y(f) df.$$

- Das aus der klassischen (deterministischen) Systemtheorie bekannte **Reziprozitätsgesetz** von Zeitdauer und Bandbreite gilt hier ebenfalls: Eine schmale Autokorrelationsfunktion entspricht einem breiten Leistungsdichtespektrum und umgekehrt.
- Als Beschreibungsgröße verwenden wir hier die **äquivalente LDS-Bandbreite** ∇f (man spricht *Nabla-f*), ähnlich definiert wie die **äquivalente AKF-Dauer** $\nabla \tau$ in Kapitel 4.4:

$$\nabla f_x = \frac{1}{\Phi_x(f=0)} \cdot \int_{-\infty}^{+\infty} \Phi_x(f) df, \quad \nabla \tau_x = \frac{1}{\varphi_x(\tau=0)} \cdot \int_{-\infty}^{+\infty} \varphi_x(\tau) d\tau.$$

- Mit diesen Definitionen gilt der folgende grundlegende Zusammenhang:

$$\nabla \tau_x \cdot \nabla f_x = 1 \quad \text{bzw.} \quad \nabla \tau_y \cdot \nabla f_y = 1.$$

Beispiel: Wir gehen wieder von der **Grafik** auf der letzten Seite aus:

- Die Kenngrößen des höherfrequenten Signals $x(t)$ sind $\nabla \tau_x = 0.33 \mu\text{s}$ und $\nabla f_x = 3 \text{ MHz}$.
- Die äquivalente AKF-Dauer des Signals $y(t)$ ist dreimal so groß: $\nabla \tau_y = 1 \mu\text{s}$.
- Die äquivalente LDS-Bandbreite beträgt somit nur mehr $\nabla f_y = \nabla f_x/3 = 1 \text{ MHz}$.

Beweis: Entsprechend den obigen Definitionen gilt:

$$\nabla \tau_x = \frac{1}{\varphi_x(\tau=0)} \cdot \int_{-\infty}^{+\infty} \varphi_x(\tau) d\tau = \frac{\Phi_x(f=0)}{\varphi_x(\tau=0)},$$

$$\nabla f_x = \frac{1}{\Phi_x(f=0)} \cdot \int_{-\infty}^{+\infty} \Phi_x(f) df = \frac{\varphi_x(\tau=0)}{\Phi_x(f=0)}.$$

\Rightarrow Das Produkt aus äquivalenter AKF-Dauer und äquivalenter LDS-Bandbreite ist gleich 1.

Ein Grenzfall des Reziprozitätsgesetzes stellt das so genannte **Weißes Rauschen** dar. Dieses beinhaltet alle Spektralanteile (bis ins Unendliche) und die äquivalente LDS-Bandbreite ∇f ist unendlich groß. Das hier angegebene Gesetz besagt dann, dass damit für die äquivalente AKF-Dauer $\nabla \tau = 0$ gelten muss; die AKF des weißen Rauschens ist diracförmig.

Mehr zu dieser Thematik finden Sie in den nachfolgenden Lernvideos, insbesondere im Teil 2:

Der AWGN-Kanal – Teil 1 (Dauer 6:00)

Der AWGN-Kanal – Teil 2 (Dauer 5:15)

Der AWGN-Kanal – Teil 3 (Dauer 6:15)

Leistungsdichtespektrum mit Gleichsignalkomponente

Wir gehen zunächst von einem gleichsignalfreien Zufallsprozess $\{x_i(t)\}$ aus. Weiterhin setzen wir voraus, dass der Prozess keinen Gleichanteil und keine periodischen Anteile beinhaltet. Dann gilt:

- Die Autokorrelationsfunktion (AKF) $\varphi_x(\tau)$ verschwindet für $\tau \rightarrow \infty$.
- Das Leistungsdichtespektrum (LDS) $\Phi_x(f)$ – berechenbar als die Fouriertransformierte von $\varphi_x(\tau)$ – ist sowohl wert- als auch zeitkontinuierlich und weist keine diskreten Anteile auf.

Wir betrachten nun einen zweiten Zufallsprozess $\{y_i(t)\}$, der sich vom Prozess $\{x_i(t)\}$ lediglich durch eine zusätzliche **Gleichsignalkomponente** unterscheidet:

$$\{y_i(t)\} = \{x_i(t) + m_y\}.$$

Die statistischen Beschreibungsgrößen des mittelwertbehafteten Zufallsprozesses $\{y_i(t)\}$ weisen dann folgende Eigenschaften auf:

- Der Grenzwert der Autokorrelationsfunktion für $\tau \rightarrow \infty$ ist nun nicht mehr Null, sondern m_y^2 . Im gesamten τ -Bereich von $-\infty$ bis $+\infty$ ist die AKF $\varphi_y(\tau)$ um m_y^2 größer als $\varphi_x(\tau)$:

$$\varphi_y(\tau) = \varphi_x(\tau) + m_y^2.$$

- Nach den elementaren Gesetzen der Fouriertransformation führt der konstante AKF-Beitrag im Leistungsdichtespektrum zu einer Diracfunktion $\delta(f)$ mit dem Gewicht m_y^2 :

$$\Phi_y(f) = \Phi_x(f) + m_y^2 \cdot \delta(f).$$

Nähere Informationen zur Diracfunktion sind im **Kapitel 2.2** des Buches „Signaldarstellung“ zu finden. Weiterhin möchten wir Sie auf das folgende Lernvideo hinweisen:

Herleitung und Visualisierung der Diracfunktion (Dauer: 2:50)

Numerische LDS-Ermittlung

Autokorrelationsfunktion und Leistungsdichtespektrum sind über die **Fouriertransformation** streng miteinander verknüpft. Dieser Zusammenhang gilt auch bei zeitdiskreter AKF-Darstellung, also für

$$A\{\varphi_x(\tau)\} = \varphi_x(\tau) \cdot \sum_{k=-\infty}^{\infty} T_A \cdot \delta(\tau - k \cdot T_A).$$

Der Übergang vom Zeit- in den Spektralbereich kann mit folgenden Schritten hergeleitet werden:

- Der Abstand T_A zweier Abtastwerte ist durch die absolute Bandbreite B_x (maximal auftretende Frequenz innerhalb des Prozesses) über das Abtasttheorem festgelegt:

$$T_A \leq \frac{1}{2B_x}.$$

- Die Fouriertransformierte der zeitdiskreten AKF ergibt ein mit $1/T_A$ periodisches LDS:

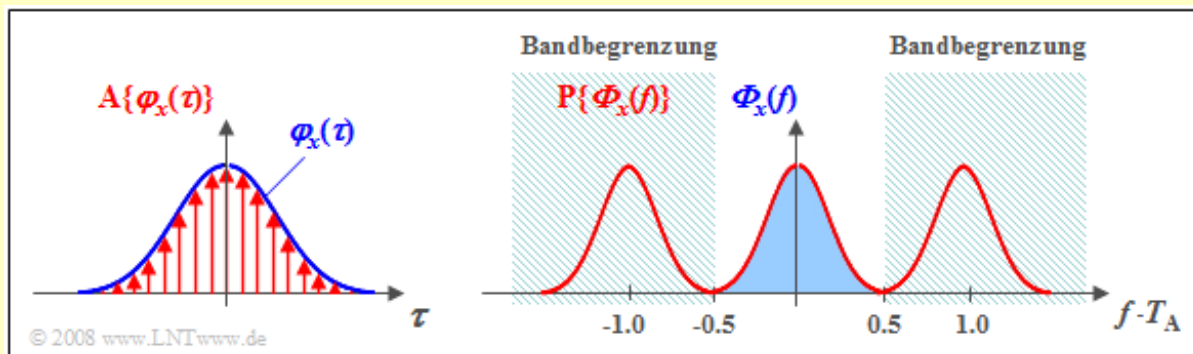
$$A\{\varphi_x(\tau)\} \circ \bullet P\{\Phi_x(f)\} = \sum_{\mu=-\infty}^{\infty} \Phi_x(f - \frac{\mu}{T_A}).$$

- Da sowohl $\varphi_x(\tau)$ als auch $\Phi_x(f)$ gerade und reelle Funktionen sind, gilt der Zusammenhang:

$$P\{\Phi_x(f)\} = T_A \cdot \varphi_x(k=0) + 2T_A \cdot \sum_{k=1}^{\infty} \varphi_x(kT_A) \cdot \cos(2\pi kT_A).$$

- Das Leistungsdichtespektrum (LDS) des zeitkontinuierlichen Prozesses erhält man aus $P\{\Phi_x(f)\}$ durch Bandbegrenzung auf den Frequenzbereich $|f| \leq 1/(2T_A)$.
- Im Zeitbereich bedeutet diese Operation eine Interpolation der einzelnen AKF-Abtastwerte mit der si-Funktion, wobei $\text{si}(x)$ für $\sin(x)/x$ steht.

Beispiel: Eine gaußförmige AKF $\varphi_x(\tau)$ wird im Abstand T_A abgetastet; das Abtasttheorem ist erfüllt. Die Fouriertransformierte der zeitdiskreten AKF $A\{\varphi_x(\tau)\}$ wird mit $P\{\Phi_x(f)\}$ bezeichnet. Dieses ist periodisch mit $1/T_A$ und dementsprechend unendlich weit ausgedehnt. In der Grafik ist $P\{\Phi_x(f)\}$ als roter Kurvenzug zu erkennen.



Das LDS $\Phi_x(f)$ des zeitkontinuierlichen Prozesses $\{x(t)\}$ erhält man durch Bandbegrenzung auf den im Bild blau hinterlegten Frequenzbereich $|f \cdot T_A| \leq 0.5$.

Genauigkeit der numerischen LDS-Berechnung (1)

Für die nachfolgende Analyse gehen wir von folgenden Annahmen aus:

- Die zeitdiskrete AKF $\varphi_x(k \cdot T_A)$ wurde aus N Abtastwerten numerisch ermittelt. Wie bereits auf der Seite **Genauigkeit der numerischen AKF-Berechnung** gezeigt wurde, sind diese Werte fehlerhaft und die Fehler korreliert, wenn N nicht hinreichend groß gewählt wurde.
- Zur Berechnung des periodischen Leistungsdichtespektrums (LDS) verwenden wir nur die AKF-Werte $\varphi_x(0), \dots, \varphi_x(K \cdot T_A)$:

$$P\{\Phi_x(f)\} = T_A \cdot \varphi_x(k=0) + 2T_A \cdot \sum_{k=1}^K \varphi_x(kT_A) \cdot \cos(2\pi kT_A).$$

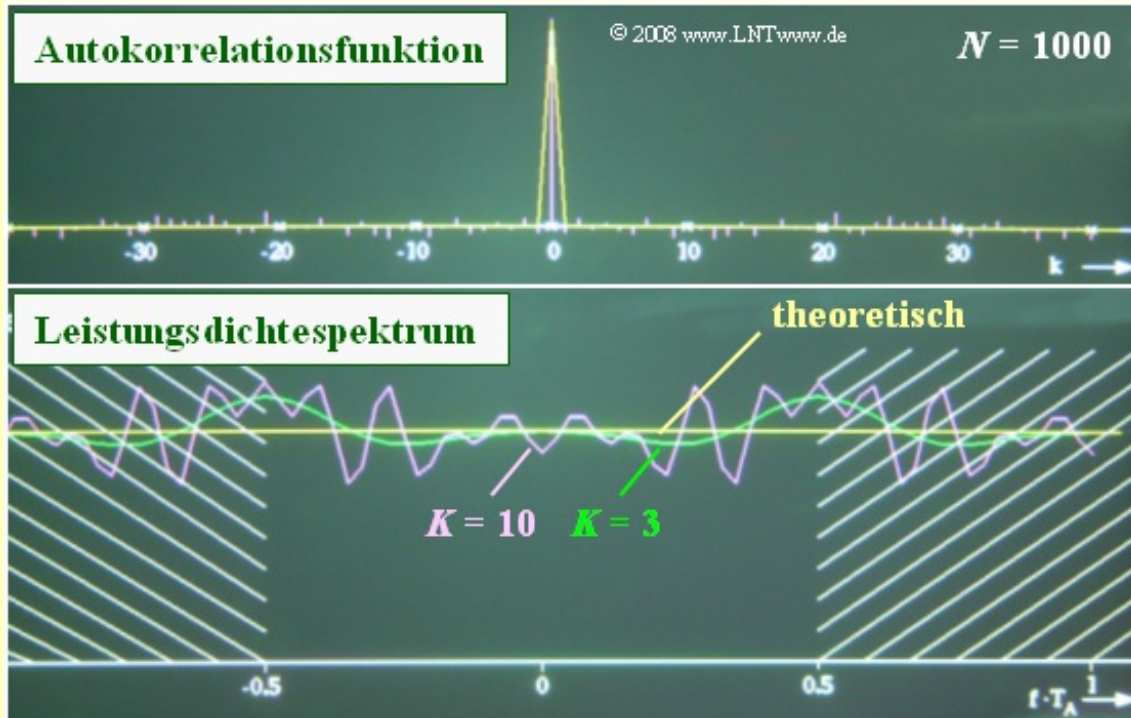
Die Genauigkeit der LDS-Berechnung wird im starken Maße durch den Parameter K bestimmt:

- Ist K zu klein gewählt, so werden die AKF-Werte $\varphi_x(k \cdot T_A)$ mit $k > K$ nicht berücksichtigt.
- Ist K zu groß gewählt, so werden auch solche AKF-Werte berücksichtigt, die eigentlich Null sein sollten und nur durch die numerische AKF-Berechnung endliche Werte besitzen.
- Diese Werte sind allerdings – bedingt durch ein zu kleines N bei der AKF-Ermittlung – nur Fehler, und beeinträchtigen die LDS-Berechnung mehr als dass sie einen brauchbaren Beitrag zum Ergebnis liefern.

Diese Aussagen werden auf der nächsten Seite anhand eines Beispiels verdeutlicht.

Genauigkeit der numerischen LDS-Berechnung (2)

Beispiel: Wir betrachten hier einen mittelwertfreien Prozess mit statistisch unabhängigen Abtastwerten, so dass nur der AKF-Wert $\varphi_x(0) = \sigma_x^2$ von 0 verschieden ist. Ermittelt man die AKF numerisch aus lediglich $N = 1000$ Abtastwerten, so erhält man auch für $k \neq 0$ endliche AKF-Werte. Das obere Bild zeigt, dass diese fehlerhaften Werte bis zu 6% des Maximalwertes betragen können.



Unten ist das numerisch ermittelte LDS dargestellt. Die gelbe Kurve zeigt den theoretischen Verlauf, der im Bereich $|f \cdot T_A| \leq 0.5$ konstant sein sollte. Die grüne und die violette Kurve verdeutlichen, wie durch $K = 3$ bzw. $K = 10$ das Ergebnis gegenüber $K = 0$ verfälscht wird.

In diesem Fall (statistisch unabhängige Zufallsgrößen) wächst der Fehler monoton mit steigendem K . Bei einer Zufallsgröße mit statistischen Bindungen gibt es dagegen jeweils einen optimalen Wert für K . Wird dieser zu klein gewählt, so werden signifikante Bindungen nicht berücksichtigt. Ein zu großer Wert führt dagegen zu Oszillationen, die nur auf fehlerhafte AKF-Werte zurückzuführen sind.

Definition der Kreuzkorrelationsfunktion

Bei vielen technischen Anwendungen interessiert man sich für ein quantitatives Maß zur Beschreibung der statistischen Verwandtschaft zwischen verschiedenen Prozessen bzw. zwischen deren Mustersignalen. Ein solches Maß ist die *Kreuzkorrelationsfunktion* (KKF), die hier unter den Voraussetzungen von *Stationarität* und *Ergodizität* hergeleitet wird.

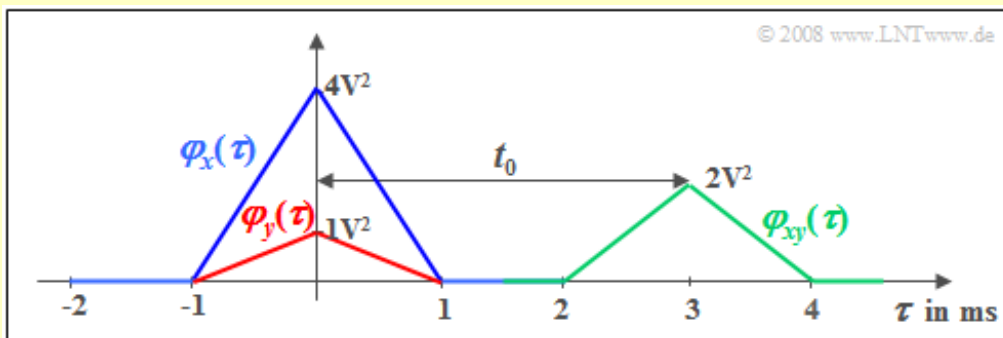
Definition: Für die **Kreuzkorrelationsfunktion** zweier stationärer und ergodischer Prozesse mit den Musterfunktionen $x(t)$ und $y(t)$ gilt:

$$\varphi_{xy}(\tau) = E[x(t) \cdot y(t + \tau)] = \lim_{T_M \rightarrow \infty} \frac{1}{T_M} \cdot \int_{-T_M/2}^{T_M/2} x(t) \cdot y(t + \tau) dt.$$

Die erste Definitionsgleichung kennzeichnet die **Erwartungswertbildung** (Scharmittelung), während die zweite Gleichung die **Zeitmittelung** über eine (möglichst große) Messdauer T_M beschreibt.

Ein Vergleich mit der **AKF-Definition** zeigt viele Gemeinsamkeiten mit dieser. Setzt man $y(t) = x(t)$, so erhält man $\varphi_{xy}(\tau) = \varphi_{xx}(\tau)$, also die Autokorrelationsfunktion, für die in Abschnitt 4.4 die vereinfachte Schreibweise $\varphi_x(\tau)$ eingeführt wurde.

Beispiel: Wir betrachten ein Zufallssignal $x(t)$ mit dreieckförmiger AKF $\varphi_x(\tau) \Rightarrow$ blaue Kurve. Diese AKF-Form ergibt sich zum Beispiel für ein Binärsignal mit gleichwahrscheinlichen bipolaren Amplitudenkoeffizienten (+1 bzw. -1) und bei rechteckförmigem Grundimpuls.



Betrachten wir dazu noch ein zweites Signal

$$y(t) = \alpha \cdot x(t - t_0),$$

das sich von $x(t)$ nur durch einen Dämpfungsfaktor α und eine Laufzeit t_0 unterscheidet, wobei dem obigen Bild $\alpha = 0.5$ und $t_0 = 3$ ms zugrunde liegen. Dieses gedämpfte und verschobene Signal besitzt die rot gezeichnete AKF

$$\varphi_y(\tau) = \alpha^2 \cdot \varphi_x(\tau).$$

Die Verschiebung um t_0 ist in der AKF nicht zu erkennen im Gegensatz zur (grün dargestellten) KKF, für die folgende Beziehung gilt:

$$\varphi_{xy}(\tau) = \alpha \cdot \varphi_x(\tau - t_0).$$

Eigenschaften der Kreuzkorrelationsfunktion

Nachfolgend sind wesentliche Eigenschaften der Kreuzkorrelationsfunktion zusammengestellt und die wichtigsten Unterschiede zur AKF herausgearbeitet.

- Die Bildung der Kreuzkorrelationsfunktion ist **nicht kommutativ**. Vielmehr gibt es stets zwei unterschiedliche Funktionen, nämlich

$$\varphi_{xy}(\tau) = E[x(t) \cdot y(t + \tau)] = \lim_{T_M \rightarrow \infty} \frac{1}{T_M} \cdot \int_{-T_M/2}^{T_M/2} x(t) \cdot y(t + \tau) dt,$$

$$\varphi_{yx}(\tau) = E[y(t) \cdot x(t + \tau)] = \lim_{T_M \rightarrow \infty} \frac{1}{T_M} \cdot \int_{-T_M/2}^{T_M/2} y(t) \cdot x(t + \tau) dt.$$

- Zwischen den beiden Funktionen besteht der Zusammenhang $\varphi_{yx}(\tau) = \varphi_{xy}(-\tau)$. Im Beispiel auf der letzten Seite hätte $\varphi_{yx}(\tau)$ sein Maximum bei $\tau = -3$ ms.
- Im Allgemeinen tritt das **KKF-Maximum** nicht bei $\tau = 0$ auf (Ausnahme: $y = \alpha \cdot x$) und dem KKF-Wert $\varphi_{xy}(\tau = 0)$ kommt keine besondere, physikalisch interpretierbare Bedeutung zu wie bei der AKF, bei der dieser Wert die Prozessleistung wiedergibt.
- Der Betrag der KKF ist nach der **Schwarzschen Ungleichung** für alle τ -Werte kleiner oder gleich dem geometrischen Mittel der beiden Signalleistungen:

$$\varphi_{xy}(\tau) \leq \sqrt{\varphi_x(\tau = 0) \cdot \varphi_y(\tau = 0)}.$$

Im Beispiel auf der letzten Seite gilt das Gleichheitszeichen:

$$\varphi_{xy}(\tau = t_0) = \sqrt{\varphi_x(\tau = 0) \cdot \varphi_y(\tau = 0)} = \alpha \cdot \varphi_x(\tau = 0).$$

- Beinhalten $x(t)$ und $y(t)$ keinen gemeinsamen periodischen Anteil, so zeigt der **Grenzwert der KKF** für $\tau \rightarrow \infty$ das Produkt der beiden Mittelwerte an:

$$\lim_{\tau \rightarrow \infty} \varphi_{xy}(\tau) = m_x \cdot m_y.$$

- Sind zwei Signale $x(t)$ und $y(t)$ **unkorreliert**, so gilt $\varphi_{xy}(\tau) \equiv 0$, das heißt, es ist $\varphi_{xy}(\tau) = 0$ für alle Werte von τ . Diese Annahme ist beispielsweise bei der gemeinsamen Betrachtung eines Nutz- und eines Störsignals in den meisten Fällen gerechtfertigt.
- Es ist jedoch stets zu beachten, dass die KKF nur die **linearen statistischen Bindungen** zwischen den Signalen $x(t)$ und $y(t)$ beinhaltet. Bindungen anderer Art – wie beispielsweise für den Fall $y(t) = x(t)^2$ – werden dagegen bei der KKF-Bildung nicht berücksichtigt.

Anwendungen der Kreuzkorrelationsfunktion

Die Anwendungen der KKF in Nachrichtensystemen sind vielfältig. Hier einige Beispiele:

- Bei **Amplitudenmodulation**, aber auch bei **BPSK-Systemen** (*Binary Phase Shift Keying*) wird zur Demodulation (Rücksetzung des Signals in den ursprünglichen Frequenzbereich) sehr häufig der so genannte Synchrondemodulator verwendet, wobei auch beim Empfänger ein Trägersignal zugesetzt werden muss, und zwar frequenz- und phasensynchron zum Sender. Bildet man die KKF zwischen dem Empfangssignal und dem empfangsseitigen Trägersignal, so lässt sich anhand der Spitze der KKF die phasensynchrone Lage zwischen den beiden Signalen erkennen, und es kann bei Auseinanderdriften nachgeregelt werden.
- Das Mehrfachzugriffsverfahren **CDMA** (*Code Division Multiple Access*) wird zum Beispiel im Mobilfunkstandard UMTS angewendet. Es erfordert eine strenge Phasensynchronität, und zwar bezüglich der zugesetzten Pseudonoise-Folgen beim Sender (*Bandspreizung*) und beim Empfänger (*Bandstauchung*). Auch dieses Synchronisationsproblem löst man meist mittels der Kreuzkorrelationsfunktion.
- Mit Hilfe der Kreuzkorrelationsfunktion kann festgestellt werden, ob ein bekanntes Signal $s(t)$ in einem verrauschten Empfangssignal $r(t) = \alpha \cdot s(t - t_0) + n(t)$ vorhanden ist oder nicht, und wenn ja, zu welchem Zeitpunkt t_0 es auftritt. Aus dem berechneten Wert für t_0 lässt sich dann beispielsweise eine Fahrgeschwindigkeit ermitteln (**Radartechnik**). Diese Aufgabenstellung kann auch mit dem so genannten Matched-Filter gelöst werden, das in Kapitel 5.4 noch eingehend beschrieben wird und das viele Gemeinsamkeiten mit der Kreuzkorrelationsfunktion aufweist.
- Beim so genannten **Korrelationsempfänger** verwendet man die KKF zur Signaldetektion. Hierbei bildet man die Kreuzkorrelation zwischen dem durch Rauschen und eventuell auch durch Verzerrungen verfälschten Empfangssignal $r(t)$ und allen möglichen Sendesignalen $s_i(t)$, wobei für den Laufindex $i = 1, \dots, I$ gelten soll. Entscheidet man N Binärsymbole gemeinsam, so ist $I = 2^N$. Man entscheidet sich dann für die Symbolfolge mit dem größten KKF-Wert und erreicht so die minimale Fehlerwahrscheinlichkeit entsprechend der *Maximum-Likelihood-Entscheidungsregel*.

Kreuzleistungsdichtespektrum

Für manche Anwendungen kann es durchaus vorteilhaft sein, die Korrelation zwischen zwei Zufallssignalen im Frequenzbereich zu beschreiben.

Definition: Die beiden **Kreuzleistungsdichtespektren** $\Phi_{xy}(f)$ und $\Phi_{yx}(f)$ ergeben sich aus den dazugehörigen Kreuzkorrelationsfunktionen durch die Fouriertransformation:

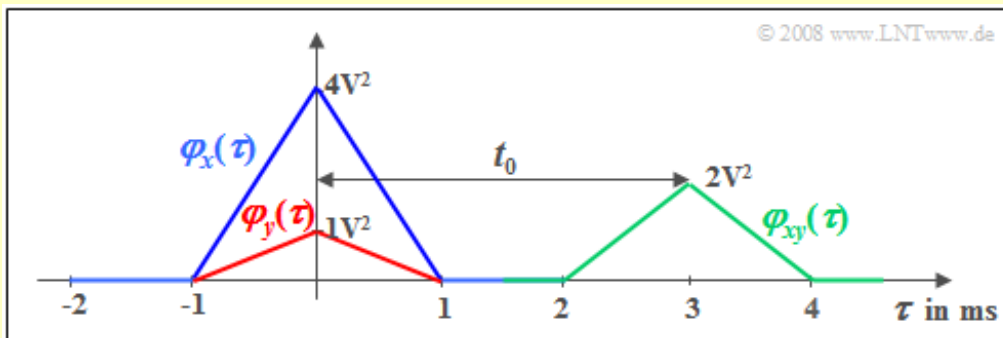
$$\Phi_{xy}(f) = \int_{-\infty}^{+\infty} \varphi_{xy}(\tau) \cdot e^{-j\pi f \tau} d\tau,$$

$$\Phi_{yx}(f) = \int_{-\infty}^{+\infty} \varphi_{yx}(\tau) \cdot e^{-j\pi f \tau} d\tau.$$

Manchmal wird hierfür auch der Begriff *spektrale Kreuzleistungsdichte* verwendet.

Es gilt hier der gleiche Zusammenhang wie zwischen einem deterministischen Signal $x(t)$ und seinem Spektrum $X(f)$ bzw. zwischen der Autokorrelationsfunktion $\varphi_x(\tau)$ eines ergodischen Prozesses $\{x_i(t)\}$ und dem dazugehörigen Leistungsdichtespektrum $\Phi_x(f)$. Ebenso beschreibt bei diesen Beispielen die **Fourierücktransformation** \Rightarrow „Zweites Fourierintegral“ den Übergang vom Spektralbereich in den Zeitbereich.

Beispiel: Wir nehmen hier Bezug zum Beispiel auf der Seite **Definition der KKF** mit den beiden „rechteckförmigen Zufallsgrößen“ $x(t)$ und $y(t) = \alpha \cdot x(t - t_0)$. Da die AKF $\varphi_x(\tau)$ dreieckförmig verläuft, ist – wie in Kapitel 4.5 beschrieben – das LDS $\Phi_x(f)$ si^2 -förmig.



Welche Aussagen können wir aus dieser Grafik für die Spektralfunktionen ableiten?

- Im zitierten Beispiel haben wir festgestellt, dass sich die Autokorrelationsfunktion $\varphi_y(\tau)$ von $\varphi_x(\tau)$ nur um den konstanten Faktor α^2 unterscheidet.
- Damit ist klar, dass das Leistungsdichtespektrum $\Phi_y(f)$ von $\Phi_x(f)$ ebenfalls nur um diesen konstanten Faktor α^2 abweicht. Beide Spektralfunktionen sind reell.
- Dagegen besitzt das Kreuzleistungsdichtespektrum einen komplexen Funktionsverlauf:

$$\Phi_{xy}(f) = \Phi_{yx}^*(f) = \alpha \cdot \Phi_x(f) \cdot e^{-j\pi f t_0}.$$

Korrelationsmatrix

Bisher wurden nur statistische Bindungen zwischen zwei (skalaren) Zufallsgrößen betrachtet. Für den allgemeineren Fall – einer Zufallsgröße mit N Dimensionen – bietet sich zweckmäßiger Weise eine Vektor- bzw. Matrixdarstellung an. Für die folgende Beschreibung wird vorausgesetzt:

- Die N -dimensionale Zufallsgröße wird als Vektor dargestellt:

$$\mathbf{x} = [x_1, x_2, \dots, x_N]^T.$$

Hierbei ist \mathbf{x} ein Spaltenvektor, was aus dem Zusatz „T“ – dies steht für „transponiert“ – des angegebenen Zeilenvektors hervorgeht.

- Die N Komponenten x_i seien jeweils eindimensionale reelle Gaußsche Zufallsgrößen.

Statistische Bindungen zwischen den N Zufallsgrößen werden durch die **Korrelationsmatrix** vollständig beschrieben:

$$\mathbf{R} = [R_{ij}] = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1N} \\ R_{21} & R_{22} & \cdots & R_{2N} \\ \dots & \dots & \dots & \dots \\ R_{N1} & R_{N2} & \cdots & R_{NN} \end{bmatrix}.$$

Die N^2 Elemente dieser $N \times N$ -Matrix geben jeweils das gemeinsame Moment erster Ordnung zwischen zwei Komponenten an:

$$R_{ij} = E[x_i \cdot x_j] = R_{ji}.$$

In Vektorschreibweise lautet somit die Korrelationsmatrix:

$$\mathbf{R} = E[\mathbf{x} \cdot \mathbf{x}^T].$$

Da \mathbf{x} ein Spaltenvektor mit N Dimensionen ist und somit der transponierte Vektor \mathbf{x}^T ein Zeilenvektor gleicher Länge, ergibt das Produkt $\mathbf{x} \cdot \mathbf{x}^T$ eine $N \times N$ -Matrix. Dagegen wäre $\mathbf{x}^T \cdot \mathbf{x}$ eine 1×1 -Matrix, also ein Skalar. Für den hier nicht weiter betrachteten Sonderfall komplexer Komponenten x_i sind auch die Matrixelemente komplex:

$$R_{ij} = E[x_i \cdot x_j^*] = R_{ji}^*.$$

Die Realteile der Korrelationsmatrix sind weiterhin symmetrisch zur Hauptdiagonalen, während sich die dazugehörigen Imaginärteile durch das Vorzeichen unterscheiden.

Kovarianzmatrix

Man kommt von der Korrelationsmatrix \mathbf{R} zur so genannten **Kovarianzmatrix**

$$\mathbf{K} = [K_{ij}] = \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1N} \\ K_{21} & K_{22} & \cdots & K_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ K_{N1} & K_{N2} & \cdots & K_{NN} \end{bmatrix},$$

wenn die Matrixelemente $K_{ij} = E[(x_i - m_i) \cdot (x_j - m_j)]$ jeweils ein **Zentralmoment** erster Ordnung angeben. Mit dem Vektor $\mathbf{m} = [m_1, m_2, \dots, m_N]^T$ kann somit auch geschrieben werden:

$$\mathbf{K} = E[(\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})^T].$$

Es soll ausdrücklich darauf hingewiesen werden, dass m_1 den Mittelwert der Komponente x_1 und m_2 den Mittelwert von x_2 bezeichnet – nicht etwa das Moment erster bzw. zweiter Ordnung.

Die Matrix \mathbf{K} zeigt bei reellen mittelwertfreien Gauß-Größen folgende weitere Eigenschaften:

- Das Element der i -ten Zeile und j -ten Spalte lautet mit den beiden Streuungen σ_i und σ_j und dem **Korrelationskoeffizienten** ρ_{ij} . Formelmäßig gilt $K_{ij} = \sigma_i \cdot \sigma_j \cdot \rho_{ij} = K_{ji}$.
- Berücksichtigt man noch die Beziehung $\rho_{ii} = 1$, so erhält man für die Kovarianzmatrix:

$$\mathbf{K} = [K_{ij}] = \begin{bmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho_{12} & \cdots & \sigma_1 \sigma_N \rho_{1N} \\ \sigma_2 \sigma_1 \rho_{21} & \sigma_2^2 & \cdots & \sigma_2 \sigma_N \rho_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_N \sigma_1 \rho_{N1} & \sigma_N \sigma_2 \rho_{N2} & \cdots & \sigma_N^2 \end{bmatrix}.$$

- Aufgrund der Beziehung $\rho_{ij} = \rho_{ji}$ ist die Kovarianzmatrix bei reellen Größen symmetrisch zur Hauptdiagonalen. Bei komplexen Größen würde dagegen $\rho_{ij} = \rho_{ji}^*$ gelten.

Beispiel: Wir betrachten drei Kovarianzmatrizen:

$$\mathbf{K}_2 = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}, \quad \mathbf{K}_3 = 4 \cdot \begin{bmatrix} 1 & 1/2 & 1/4 \\ 1/2 & 1 & 3/4 \\ 1/4 & 3/4 & 1 \end{bmatrix}, \quad \mathbf{K}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 16 \end{bmatrix}.$$

- \mathbf{K}_2 beschreibt eine 2D-Zufallsgröße, wobei der Korrelationskoeffizient ρ zwischen den zwei Komponenten -0.5 beträgt und beide Komponenten die Streuung $\sigma = 1$ aufweisen.
- Bei der 3D-Zufallsgröße gemäß \mathbf{K}_3 haben alle Komponenten die gleiche Streuung $\sigma = 2$. Die stärksten Bindungen bestehen zwischen x_2 und x_3 ; wobei $\rho_{23} = 3/4$ gilt.
- Die vier Komponenten der durch \mathbf{K}_4 gekennzeichneten Zufallsgröße sind unkorreliert, bei Gaußscher WDF auch statistisch unabhängig. Die Streuungen sind $\sigma_i = i$ für $i = 1, \dots, 4$.

Zusammenhang zwischen Kovarianzmatrix und WDF

Die *Wahrscheinlichkeitsdichtefunktion* einer N -dimensionalen Gaußschen Zufallsgröße \mathbf{x} lautet:

$$\mathbf{f}_{\mathbf{x}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N \cdot |\mathbf{K}|}} \cdot \exp \left[-\frac{1}{2} \cdot (\mathbf{x} - \mathbf{m})^T \cdot \mathbf{K}^{-1} \cdot (\mathbf{x} - \mathbf{m}) \right].$$

Hierbei bezeichnen:

- \mathbf{x} den Spaltenvektor der betrachteten N -dimensionalen Zufallsgröße,
- \mathbf{m} den Spaltenvektor der zugehörigen Mittelwerte,
- $|\mathbf{K}|$ die Determinante der $N \times N$ -Kovarianzmatrix \mathbf{K} – eine skalare Größe,
- \mathbf{K}^{-1} die Inverse von \mathbf{K} ; diese ist ebenfalls eine $N \times N$ -Matrix.

Die Multiplikation des Zeilenvektors $(\mathbf{x} - \mathbf{m})^T$, der Matrix \mathbf{K}^{-1} und des Spaltenvektors $(\mathbf{x} - \mathbf{m})$ ergibt im Argument der Exponentialfunktion erwartungsgemäß ein Skalar.

Beispiel: Wir betrachten wie im **Beispiel** auf der letzten Seite wieder eine 4D-Zufallsgröße \mathbf{x} , deren Kovarianzmatrix nur auf der Hauptdiagonalen besetzt ist:

$$\mathbf{K} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}.$$

Deren Determinante ist $|\mathbf{K}| = \sigma_1^2 \cdot \sigma_2^2 \cdot \sigma_3^2 \cdot \sigma_4^2$. Die inverse Kovarianzmatrix ergibt sich zu:

$$\mathbf{K}^{-1} \cdot \mathbf{K} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{K}^{-1} = \begin{bmatrix} \sigma_1^{-2} & 0 & 0 & 0 \\ 0 & \sigma_2^{-2} & 0 & 0 \\ 0 & 0 & \sigma_3^{-2} & 0 \\ 0 & 0 & 0 & \sigma_4^{-2} \end{bmatrix}.$$

Für mittelwertfreie Größen ($\mathbf{m} = \mathbf{0}$) lautet somit die WDF:

$$\mathbf{f}_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(2\pi)^2 \cdot \sigma_1 \cdot \sigma_2 \cdot \sigma_3 \cdot \sigma_4} \cdot \exp \left[-\left(\frac{x_1^2}{2\sigma_1^2} + \frac{x_2^2}{2\sigma_2^2} + \frac{x_3^2}{2\sigma_3^2} + \frac{x_4^2}{2\sigma_4^2} \right) \right].$$

Ein Vergleich mit **Kapitel 4.2** zeigt, dass es sich um eine 4D-Zufallsgröße mit statistisch unabhängigen und unkorrelierten Komponenten handelt, da folgende Bedingung erfüllt ist:

$$\mathbf{f}_{\mathbf{x}}(\mathbf{x}) = \mathbf{f}_{x_1}(x_1) \cdot \mathbf{f}_{x_2}(x_2) \cdot \mathbf{f}_{x_3}(x_3) \cdot \mathbf{f}_{x_4}(x_4).$$

Der Fall korrelierter Komponenten wird in **Aufgaben zu diesem Kapitel** eingehend behandelt.

Die folgenden Links verweisen auf Seiten mit Grundlagen der Matrizenrechnung am Kapitelende:

Determinante einer Matrix

Inverse einer Matrix

Eigenwerte und Eigenvektoren (1)

Wir gehen weiter von einer $N \times N$ -Kovarianzmatrix \mathbf{K} aus. Hieraus lassen sich die N **Eigenwerte** – im Folgenden mit $\lambda_1 \dots \lambda_N$ bezeichnet – wie folgt berechnen:

$$|\mathbf{K} - \lambda \cdot \mathbf{E}| = 0.$$

\mathbf{E} ist die Einheits-Diagonalmatrix der Dimension N .

Beispiel: Ausgehend von einer 2×2 -Matrix \mathbf{K} mit $K_{11} = K_{22} = 1$ und $K_{12} = K_{21} = 0.8$ erhält man als Bestimmungsgleichung:

$$\det \begin{bmatrix} 1 - \lambda & 0.8 \\ 0.8 & 1 - \lambda \end{bmatrix} = 0 \quad \Rightarrow \quad (1 - \lambda)^2 - 0.64 = 0.$$

Die beiden Eigenwerte sind somit $\lambda_1 = 1.8$ und $\lambda_2 = 0.2$.

Mit den so ermittelten Eigenwerten λ_i ($i = 1, \dots, N$) kann man die dazugehörigen **Eigenvektoren** ξ_i berechnen. Die N vektoriellen Bestimmungsgleichungen lauten dabei:

$$(\mathbf{K} - \lambda_i \cdot \mathbf{E}) \cdot \xi_i = 0 \quad (i = 1, \dots, N).$$

Beispiel: In Fortsetzung obiger Rechnung ergeben sich die beiden folgenden Eigenvektoren:

$$\begin{bmatrix} 1 - 1.8 & 0.8 \\ 0.8 & 1 - 1.8 \end{bmatrix} \cdot \xi_1 = 0 \quad \Rightarrow \quad \xi_1 = \text{const.} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$
$$\begin{bmatrix} 1 - 0.2 & 0.8 \\ 0.8 & 1 - 0.2 \end{bmatrix} \cdot \xi_2 = 0 \quad \Rightarrow \quad \xi_2 = \text{const.} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Bringt man die Eigenvektoren in die so genannte Orthonormalform (jeweils mit Betrag 1), so lauten sie:

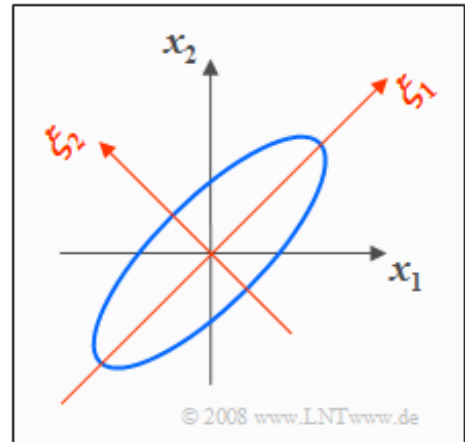
$$\xi_1 = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \xi_2 = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Eigenwerte und Eigenvektoren (2)

Abschließend soll diskutiert werden, wie Eigenwert und Eigenvektor in der Informationstechnik genutzt werden können, beispielsweise zum Zwecke der Datenreduktion.

Wir gehen von den Parameterwerten des soeben betrachteten Beispiels aus.

- Mit $\sigma_1 = \sigma_2 = 1$ und $\rho = 0.8$ ergibt sich die nachfolgend skizzierte 2D-WDF mit elliptischen Höhenlinien.
- Die Ellipsenhauptachse liegt hier wegen $\sigma_1 = \sigma_2$ unter einem Winkel von 45 Grad.



In der Grafik ist zusätzlich das (ξ_1, ξ_2) -Koordinatensystem eingezeichnet, das durch die Eigenvektoren ξ_1 und ξ_2 der Korrelationsmatrix aufgespannt wird. Die Eigenwerte $\lambda_1 = 1.8$ und $\lambda_2 = 0.2$ geben die Varianzen bezüglich des neuen Koordinatensystems an. Die Streuungen sind somit $\sigma_1 = 1.8^{0.5} \approx 1.341$ und $\sigma_2 = 0.2^{0.5} \approx 0.447$.

Beispiel: Soll eine 2D-Zufallsgröße \mathbf{x} in seinen beiden Dimensionen x_1 und x_2 im Bereich zwischen -5σ und $+5\sigma$ im Abstand $\Delta_x = 0.01$ quantisiert werden, so gibt es 10^6 mögliche Quantisierungswerte ($\sigma_1 = \sigma_2 = \sigma = 1$ vorausgesetzt).

Dagegen ist die Anzahl der möglichen Quantisierungswerte bei der gedrehten Zufallsgröße ξ um den Faktor $1.341 \cdot 0.447 \approx 0.6$ geringer. Das bedeutet: Allein durch die Drehung des Koordinatensystems um $45^\circ \Rightarrow$ Transformation der zweidimensionalen Zufallsgröße wurde eine Datenreduktion um 40% erreicht.

Die Ausrichtung entsprechend den Hauptdiagonalen wurde für den zweidimensionalen Fall bereits auf der Seite **Drehung des Koordinatensystems** im Kapitel 4.2 behandelt, und zwar basierend auf geometrischen und trigonometrischen Überlegungen. Die Lösung des Problems mit Eigenwert und Eigenvektor ist äußerst elegant und zudem problemlos auf beliebig große Dimensionen N erweiterbar.

Grundlagen der Matrizenrechnung: Determinante einer Matrix

Wir betrachten die beiden quadratischen Matrizen mit Dimension $N = 2$ bzw. $N = 3$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}.$$

Die beiden **Determinanten** dieser Matrizen lauten:

$$|\mathbf{A}| = a_{11} \cdot a_{22} - a_{12} \cdot a_{21},$$

$$|\mathbf{B}| = b_{11} \cdot b_{22} \cdot b_{33} + b_{12} \cdot b_{23} \cdot b_{31} + b_{13} \cdot b_{21} \cdot b_{32} - \\ - b_{11} \cdot b_{23} \cdot b_{32} - b_{12} \cdot b_{21} \cdot b_{33} - b_{13} \cdot b_{22} \cdot b_{31}.$$

Bitte beachten Sie:

- Die Determinante der Matrix \mathbf{A} lässt sich geometrisch als die Fläche des durch die beiden Zeilenvektoren (a_{11}, a_{12}) und (a_{21}, a_{22}) aufgespannten Parallelogramms interpretieren.
- Die Fläche des durch die beiden Spaltenvektoren $(a_{11}, a_{21})^T$ und $(a_{12}, a_{22})^T$ festgelegten Parallelogramms ist ebenfalls $|\mathbf{A}|$.
- Dagegen ist die Determinante der Matrix \mathbf{B} bei analoger geometrischer Interpretation als Volumen zu verstehen.

Für $N > 2$ ist es möglich, sogenannte Unterdeterminanten zu bilden. Die **Unterdeterminante** einer $N \times N$ -Matrix bezüglich der Stelle i, j ist die Determinante der $(N-1) \times (N-1)$ -Matrix, die sich ergibt, wenn man die i -te Zeile und die j -te Spalte streicht. Als **Kofaktor** bezeichnet man dann den Wert der Unterdeterminante gewichtet mit dem Vorzeichen $(-1)^{i+j}$.

Beispiel: Ausgehend von der 3×3 -Matrix \mathbf{B} lauten die Kofaktoren der zweiten Zeile:

$$B_{21} = -(b_{12} \cdot b_{23} - b_{13} \cdot b_{32}) \quad \text{da } i + j = 3,$$

$$B_{22} = +(b_{11} \cdot b_{23} - b_{13} \cdot b_{31}) \quad \text{da } i + j = 4,$$

$$B_{23} = -(b_{11} \cdot b_{32} - b_{12} \cdot b_{31}) \quad \text{da } i + j = 5.$$

Die Determinante von \mathbf{B} ergibt sich mit diesen Kofaktoren zu:

$$|\mathbf{B}| = b_{21} \cdot B_{21} + b_{22} \cdot B_{22} + b_{23} \cdot B_{23}.$$

Die Determinante wurde hier nach der zweiten Zeile entwickelt. Entwickelt man \mathbf{B} nach einer anderen Zeile oder Spalte, so ergibt sich für $|\mathbf{B}|$ der gleiche Zahlenwert.

Grundlagen der Matrizenrechnung: Inverse einer Matrix

Häufig benötigt man die Inverse \mathbf{M}^{-1} der quadratischen Matrix \mathbf{M} . Die inverse Matrix \mathbf{M}^{-1} besitzt die gleiche Dimension N wie \mathbf{M} und ist wie folgt definiert, wobei \mathbf{E} die Einheitsmatrix (Diagonalmatrix) bezeichnet:

$$\mathbf{M}^{-1} \cdot \mathbf{M} = \mathbf{E} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Die Inverse der 2×2 -Matrix \mathbf{A} lautet demnach:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Hierbei gibt $|\mathbf{A}| = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$ die **Determinante** an. Entsprechend gilt für $N = 3$:

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}^{-1} = \frac{1}{|\mathbf{B}|} \cdot \begin{bmatrix} B_{11} & B_{21} & B_{31} \\ B_{12} & B_{22} & B_{32} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}.$$

Die Determinante $|\mathbf{B}|$ einer 3×3 -Matrix wurde auf der letzten Seite angegeben, ebenso wie die Vorschrift zur Berechnung der Kofaktoren B_{ij} . Diese beschreiben die Unterdeterminanten von \mathbf{B} , gewichtet mit den Positionsvorzeichen $(-1)^{i+j}$. Zu beachten ist die Vertauschung der Zeilen und Spalten bei der Inversen.