4 Classical Coherence Theory

This chapter is based largely on Wolf, *Introduction to the theory of coherence and polarization of light* [**?**]. Until now, we have not been concerned with the nature of the light field itself. Instead, we studied the behavior of light waves in different optical systems. In many cases the wave fields have been monochromatic; sometimes they have consisted of plane waves, sometimes we have considered point sources. In all cases, we have *implicitly* assumed that the fields possessed the property of complete temporal and spatial coherence: the phase of a wave field at point $\mathbf{r}_0$ at time t_0 has determined the phase for all other times $t > t_0$ at all other points $\mathbf{r}$.

However, not even a single free atom radiates strictly monochromatic light:

$\Rightarrow$ Life time of excited states $\tau_c \sim 10^{-8}$sec corresponds to length $l_c = c\tau_c \sim 3m$

$\Rightarrow$ Frequency uncertainty $\triangle\omega \sim \dfrac{1}{\tau_c}$ due to this finite life time (or more precisely finite transition time between atomic energy levels).

$\Rightarrow$ For atoms interacting with their environments the so called coherence time τ_c can be much shorter and so will be the so-called coherence length l_c

In addition, a common light source consists of a very large number of emitters which radiate more or less independently.

$\Rightarrow$ For fixed time t_0 and close to the light source this leads to reduced correlations between the phases for the wave field at different locations because different atoms radiate at more or less random times (this is minimized in a Laser).

Nevertheless, interference effects (for which coherence is obviously necessary) have been observed well before the invention of the Laser. Why? One reason will be the so-called Van Cittert-Zernicke theorem which we will discuss below. First, however, we have to set up a quantitative description of coherence.

4.1 Elementary Coherence Phenomena

We consider a scalar monochromatic wave field

$$U(t) = a\cos(\phi - \omega t). \tag{4.1}$$

In view of the discussion above, we have to replace this field with a so-called quasi-monochromatic field that is characterized by the following quantities:

(Narrow) Frequency bandwidth: $\triangle\omega$

Mean frequency: $\bar{\omega}$

Quasi-monochromaticity means

$$\frac{\triangle\omega}{\bar{\omega}} \ll 1 \tag{4.2}$$

In essence, this leads to the following wave field

$$\Rightarrow U(t) = a(t)\cos(\phi(t) - \bar{\omega}t), \tag{4.3}$$

where the terms $a(t)$ and $\phi(t)$ may fluctuate randomly (see Fig. 4.1).

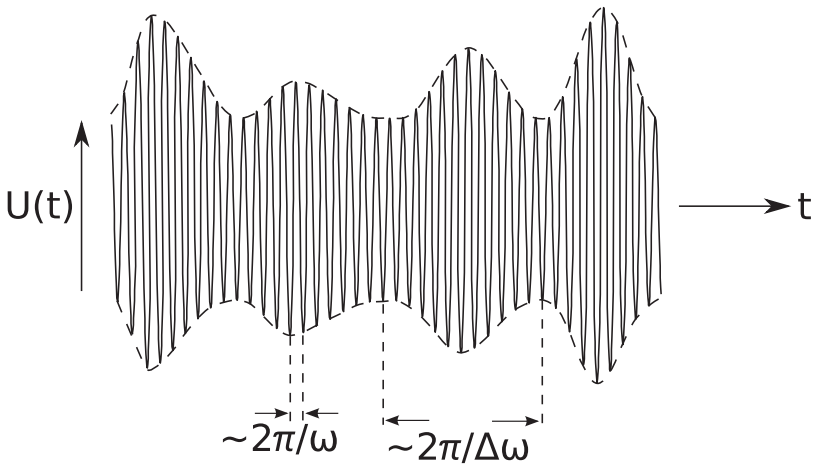

Figure 4.1: Illustration of the light fluctuations $U(t)$ of quasi-monochromatic light at a fixed position $\mathbf{r}_0$ in space.

For quasi-monochromatic light $a(t)$ and $\phi(t)$ will vary slowly over any time interval $\Delta t \ll \dfrac{2\pi}{\Delta\omega}$ (that is short compared with the reciprocal bandwidth).

Note: Here and throughout this lecture, we implicitly assume that we are dealing with a wave field whose statistical behavior which causes the fluctuations does not change over time. In other words, we assume that the wave field is statistically stationary.

We now consider at point P the interference of two quasi-monochromatic waves that emanate from two pin holes p_1, p_2, so

$$U(t) = a_1(t)\cos(\phi_1(t) - \bar{\omega}t) + a_2(t)\cos(\phi_2(t) - \bar{\omega}t + \delta). \tag{4.4}$$

Here, δ is the phase difference originating from the different propagation distances from p_1 and p_2 to P.

This gives the instantaneous intensity:

$$I(t) = U^2(t) = I_1(t) + I_2(t) + I_{12}(t) \tag{4.5}$$

with

$$\begin{aligned} I_1(t) &= a_1^2(t)\cos^2(\phi_1(t) - \bar{\omega}t), && (4.6)\\ I_2(t) &= a_2^2(t)\cos^2(\phi_2(t) - \bar{\omega}t + \delta), && (4.7)\\ I_{12}(t) &= a_1(t)a_2(t)\big[\cos\big(\phi_1(t) + \phi_2(t) - 2\bar{\omega}t + \delta\big) && (4.8)\\ &\qquad + \cos(\phi_1(t) - \phi_2(t) - \delta)\big]. \end{aligned}$$

However, at optical frequencies the instantaneous intensity cannot be measured. Instead only the average value over a certain time interval $[-T, T]$ with $T \gg \dfrac{1}{\Delta\omega}$ can be measured. We use $T \to \infty$ (justification via ergodicity hypotheses below).

$$\langle I\rangle_t = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \mathrm{d}t\; I(t). \tag{4.9}$$

Here, we have introduced the abbreviation $\langle I\rangle_t$ for the time average.

The time average becomes

$$\langle I\rangle_t = \langle I_1\rangle_t + \langle I_2\rangle_t + \langle I_{12}\rangle_t, \tag{4.10}$$

and for the case $a_1(t) = a_2(t) = a = \text{constant}$, the above equation becomes

$$\langle I_1\rangle_t = \langle I_2\rangle_t = \frac{1}{2}a^2, \tag{4.11}$$

$$\langle I_{12}\rangle_t = a^2\langle\cos(\phi_1(t) - \phi_2(t) - \delta)\rangle_t. \tag{4.12}$$

Here, $\langle I_{12}\rangle_t$ is the interference term, that may be present even if the phases $\phi_1(t)$ and $\phi_2(t)$ are not constant, or even if they fluctuate randomly, i.e.: Even the superposition of incoherent light waves can produce typical interference patterns.

Note:

- To obtain interference, light need not be monochromatic.
- For (nearly) constant phases $\phi_1(t)$, $\phi_2(t)$ the position of the observer P matters.
- These observer-dependent interference phenomena are sometimes called speckles.

4.1.1 Measuring Temporal (Longitudinal) Coherence

Temporal coherence is measured via a Michelson interferometer (see Fig. 4.2). The two arms of the interferometer lead to a relative path delay $c\Delta t$ for different spectral components. The fringe contrast depends on Δt because fringe patterns for different spectral components may overlap. The corresponding coherence length of the light is

$$\Delta l_c = \frac{2\pi c}{\Delta\omega} = \frac{\bar{\lambda}}{\Delta\lambda}\cdot\bar{\lambda} \tag{4.13}$$

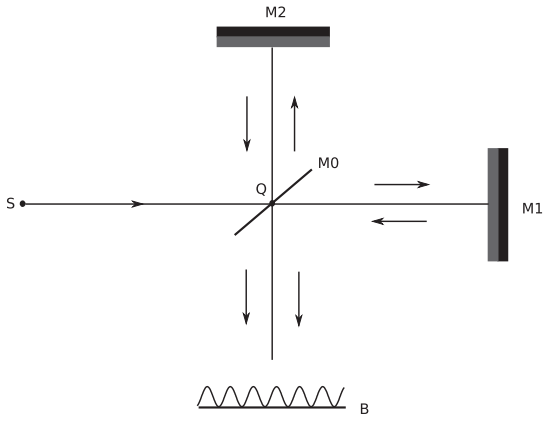

Figure 4.2: Michelson interferometer for the measurement of temporal coherence. Here, M_1 and M_2 are mirrors, while M_0 denotes a beam-splitter.

If the relative delay $c\Delta t$ is

$< \Delta l_c$, then interference fringes are visible

$> \Delta l_c$, then interference fringes are not visible

Thus, proper design of the Michelson interferometer becomes an important issue. For illustration, we list below some typical values for incandescent lamps and well-stabilized lasers.

Data	Thermal light(incandescent lamp)	Well stabilized Laser
$\Delta\omega$	$\sim 10^8\,\mathrm{s}^{-1}$	$\sim 10^4\,\mathrm{s}^{-1}$
Δt	$\sim 10^{-8}\,\mathrm{s}$	$\sim 10^{-4}\,\mathrm{s}$
Δl	$\sim 19\,\mathrm{m}$	$\sim 190\,\mathrm{km}$

An extreme Michelson interferometer is the LIGO gravitational wave detector.

4.1.2 Measuring Spatial (Transverse) Coherence

Spatial coherence is measured via Young's double slit experiment (see Fig. 4.3). We have a source with spatial extent Δx. In order to observe interference, the angle $\Delta\theta$ should obey

$$\Delta X \cdot \Delta\theta < \bar{\lambda}. \tag{4.14}$$

We further denote the area of the source with $S = (\Delta x)^2$. The coherence area is then given by

$$\Delta A = \frac{\bar{\lambda}^2}{\Delta\Omega'}, \tag{4.15}$$

with the solid angle

$$\Delta\Omega' = \frac{S}{R^2}. \tag{4.16}$$

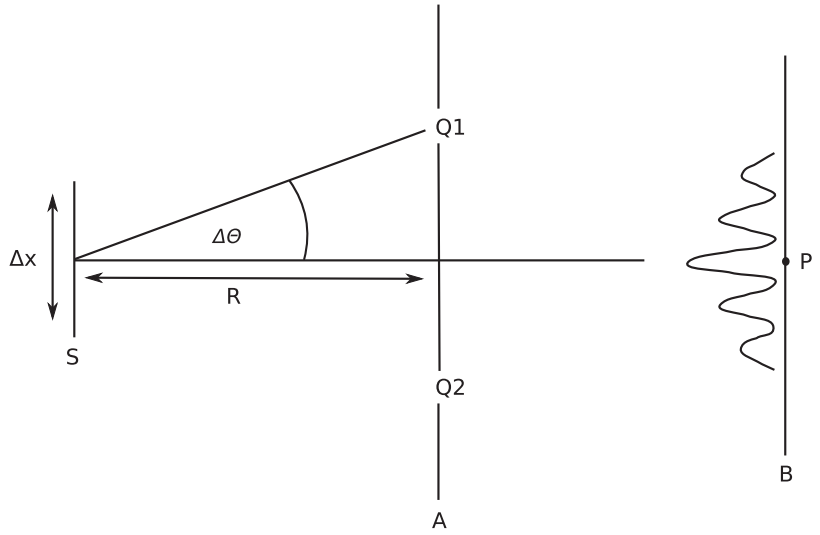

Figure 4.3: Illustration of Young's double-slit experiment, which can be used to measure spatial coherence.

Example: $S = 1\,\mathrm{mm}^2$, $\bar{\lambda} = 500\,\mathrm{nm}$, $R = 2\,\mathrm{m}$, then

Coherence area	Thermal source	Sun (500 nm filter)	Betelgeuse (500 nm filter)
ΔA	$1\,\mathrm{mm}^2$	$3.68 \times 10^{-3}\,\mathrm{mm}^2$	$6\,\mathrm{m}^2$

The apparent discrepancy between the sun and Betelgeuse (a star in the constellation Orion) is explained via the van Cittert-Zernicke theorem (see below).

Sometimes, temporal and spatial coherence are combined into what is known as a coherence volume (see Fig. 4.4).

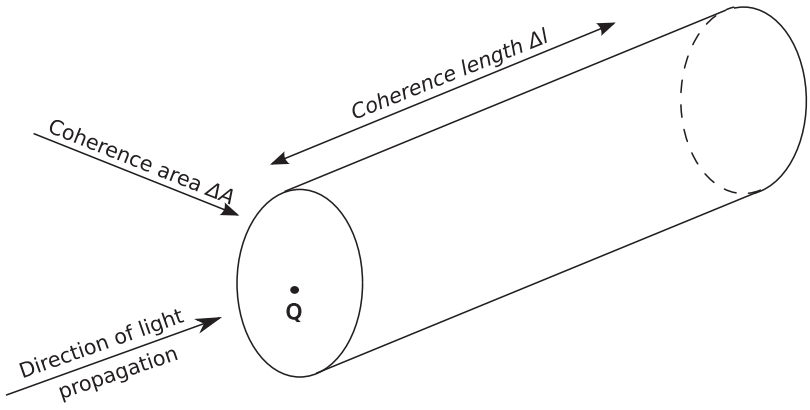

Figure 4.4: The coherence volume as a product of the coherence area ΔA and the coherence length Δl.

4.2 Mathematical Description

Theory of Random (Stochastic) Process

In this section, X(t) is randomly fluctuating field component at time t observed at some fixed point $\mathbf{r}_0$ in space. A particular realization/observation of this random variable is denoted by ${}^kX(t)$.

The exact behavior $U(t)$ cannot be predicted. Suppose that $U(t)$ is measured in a series of very similar experiments. ${}^1X(t)$, ${}^2X(t)$, ${}^3X(t)$, ..., are the outcomes of such experiments (see Fig. 4.5).

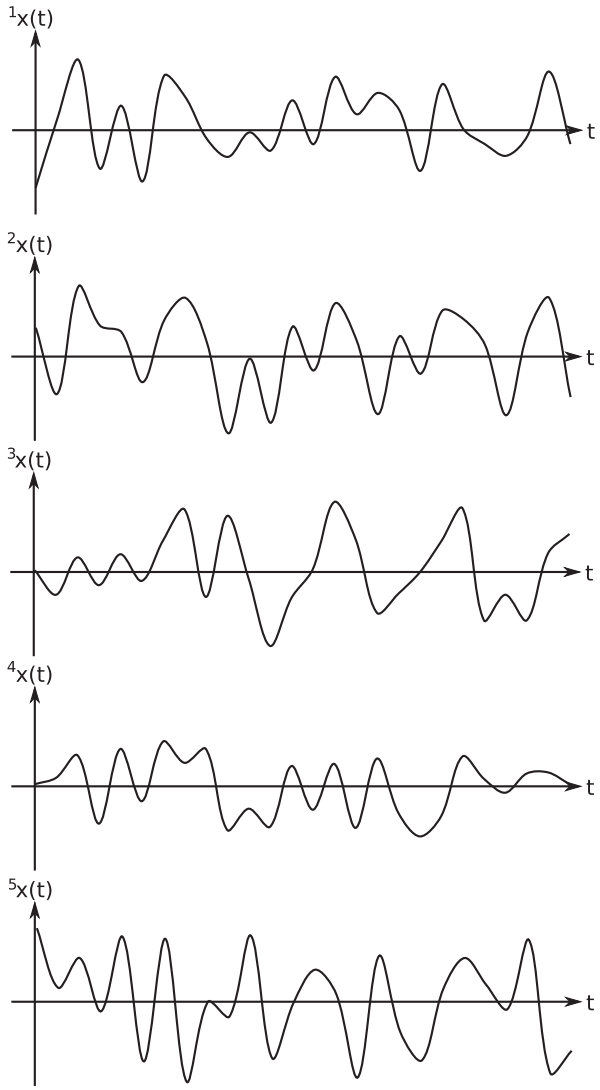

Figure 4.5: An ensemble of different realizations ${}^kX(t)$ of a random variable $X(t)$ describing a randomly fluctuating field.

This set is called an ensemble of realizations (sample functions) of the random function

(or random process) $X(t)$. In order to characterize and quantify the "randomness" of these fields, we need the following time averaged quantities:

$$\langle {}^kX(t)\rangle_t = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \mathrm{d}t\ {}^kX(t), \tag{4.17}$$

$$\langle F({}^kX(t))\rangle_t = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \mathrm{d}t\ F({}^kX(t)). \tag{4.18}$$

Here, $F(x)$ stands for any physical quantity that depends on the fields, such as energy density etc.

One can also define ensemble averages (or expectation values) $\langle X(t)\rangle_e$ via

$$\langle X(t)\rangle_e = \lim_{N\to\infty} \frac{1}{N}\sum_{K=1}^{N} \quad {}^kX(t). \tag{4.19}$$

This expression can also be formulated in the *continuum limit* as follows: Let $p_1(x,t)\mathrm{d}x$ be the probability that the random variable will take on a value in the range $(x, x+dx)$ at time t. Thus, $p_1(x,t) = \dfrac{n}{N}$, where n is the number of realizations ${}^kX(t)$ that at time t have a value in the range $(x, x+dx)$ and N is the total number of realizations observed. Then, we can define ensemble averages via

$$\langle X(t)\rangle_e = \int \mathrm{d}x\ xp_1(x,t), \tag{4.20}$$

$$\langle F(X(t))\rangle_e = \int \mathrm{d}x\ F(x)p_1(x,t). \tag{4.21}$$

Examples: For a simple incandescent lamp (ordinary light bulb with a wire), one can show that

$$p_1(U) = \frac{1}{\sqrt{2\pi\langle I\rangle}}\exp\Big(\frac{-U^2}{2\langle I\rangle}\Big), \tag{4.22}$$

where $\langle I\rangle = \langle U^2\rangle$ is the average intensity.

Similarly, simple laser theory gives for a laser

$$p_1(U) = \begin{cases} \dfrac{1}{\pi\sqrt{\langle I\rangle - U^2}}, & \text{if } |U| < \sqrt{\langle I\rangle} \\ 0, & \text{if } |U| < \sqrt{\langle I\rangle}, \end{cases} \tag{4.23}$$

where $\langle I\rangle$ is the stabilized intensity of the laser output (see Fig. 4.6).

However, to fully characterize a random process one requires a sequence of (joint) probability densities $p_1(x_1,t_1)$, $p_2(x_1,x_2;t_1,t_2)$, ... Here, $p_2dx_1dx_2$ is the probability that the variable x will take on a value in the range (x_1, x_1+dx_1) at time t_1 *and* a value in the range (x_2, x_2+dx_2) at time t_2. The corresponding averages are

$$\langle X(t_1)X(t_2)\rangle_e = \iint \mathrm{d}x_1\mathrm{d}x_2\ x_1x_2p_2(x_1,x_2;t_1,t_2). \tag{4.24}$$

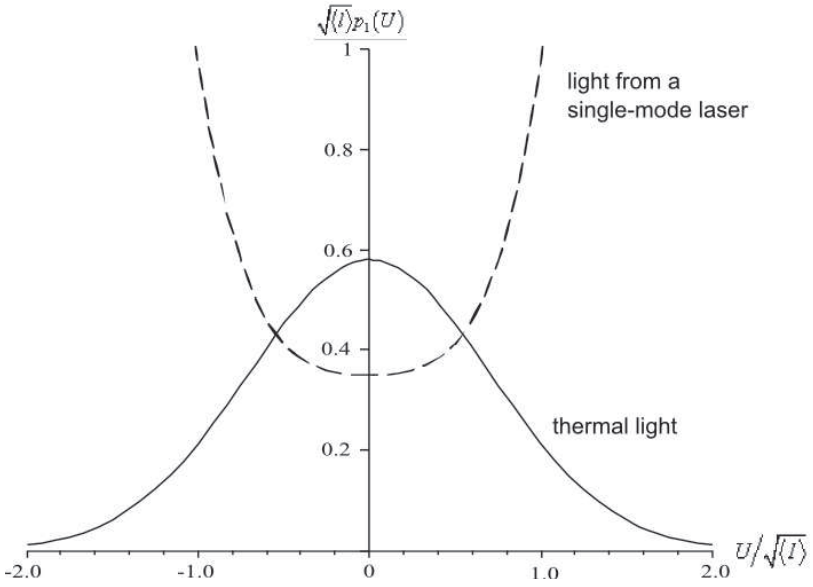

Figure 4.6: Probability density $p_1(U)$ of the field variable U for light from a thermal source and from a single-mode laser.

Furthermore, the above constructions can all be generalized to complex valued processes, for instance

$$z = x + \mathrm{i}y \in \mathbb{C}, \quad x, y \in \mathbb{R}, \tag{4.25}$$

$$Z(t) = X(t) + \mathrm{i}Y(t) \in \mathbb{C}, \quad X(t), Y(t) \in \mathbb{R}, \tag{4.26}$$

$$\langle Z^*(t_1)Z(t_2)\rangle_e = \iint \underbrace{\mathrm{d}^2 z_1}_{\mathrm{d}x_1\mathrm{d}y_1} \underbrace{\mathrm{d}^2 z_2}_{\mathrm{d}x_2\mathrm{d}y_2}\; z_1^* z_2 p_2(z_1, z_2; t_1, t_2). \tag{4.27}$$

Above, we have stated that we assume a *statistically stationary* situation without quantifying this statement. Now, we can remedy this shortcoming and define a *stationary random process* as

$$p_n(z_1, z_2, ..., z_n; t_1 + \tau, t_2 + \tau, ..., t_n + \tau) := p_n(z_1, z_2, ..., z_n; t_1, t_2, ..., t_n) \quad \forall\tau \tag{4.28}$$

$$\equiv p_n(z_1, z_2, ..., z_n). \tag{4.29}$$

This means: A random process is called stationary, if the probability distributions governing its random variables do actually not depend on time explicitly.

4.2.1 Ergodicity

A statistically stationary process is said to be *ergodic* if the time average of any deterministic function $F(\xi)$ of a typical realization $\xi = {}^kX(t)$ of the random process is equal to the corresponding ensemble average

$$\langle F[{}^kX(t)]\rangle_t = \langle F[X(t)]\rangle_e. \tag{4.30}$$

From now on: We assume to have only ergodic processes and drop the distinction between time and ensemble average, i.e., we drop the indices $\langle \quad \rangle_t, \langle \quad \rangle_e$ (see Fig. 4.7).

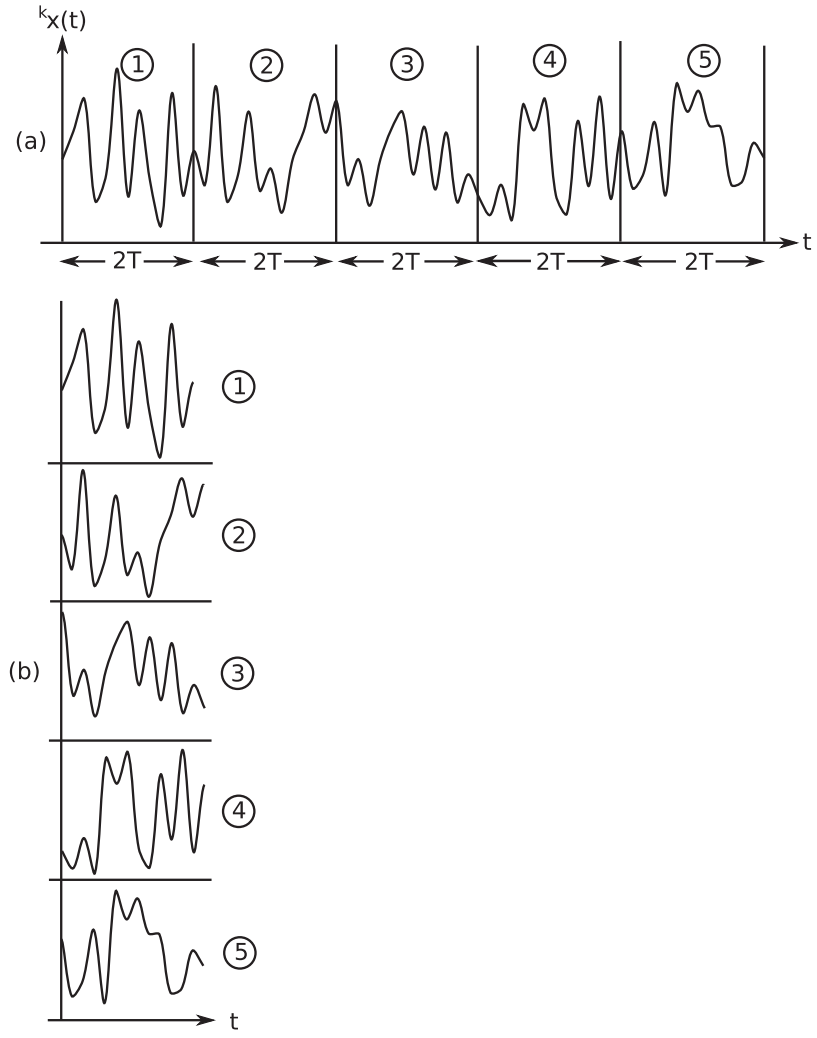

Figure 4.7: Ergodicity means that the statistical information about the ensemble of a stationary random process $X(t)$ is contained in a single sample ${}^kX(t)$. Dividing this sample into many slices in time (a), provides a valid ensemble of realizations (b).

4.2.2 Complex Representation of a Real Signal

Let $U(t)$ be a real real and fluctuating field. It can be represented by its Fourier transform $U(\omega)$ as

$$U(t) = \int_{-\infty}^{\infty} \mathrm{d}\omega \; U(\omega) \mathrm{e}^{-\mathrm{i}\omega t}, \tag{4.31}$$

with $U(-\omega) = U^*(\omega)$, because $U(t)$ is real.

Clearly, negative frequencies do not provide any new information so that we instead can as well consider the complex function

$$V(t) := \int_0^\infty \mathrm{d}\omega\, U(\omega)\mathrm{e}^{-\mathrm{i}\omega t}. \tag{4.32}$$

$V(t)$ is called the complex analytic signal (as a function of a complex variable t it is analytic in the lower part of the complex t plane) associated with the process $U(t)$.

4.2.3 Auto- and Cross-Correlation Functions

We define the following expectation values associated with a real random process $X(t)$:

Mean (value): $m := \langle X(t)\rangle$.

Auto-correlation: $R(t_1, t_2) := \langle X(t_1)X(t_2)\rangle$.

For a stationary process, with $t_2 = t_1 + \tau$ the auto-correlation actually simplifies to a function of one variable τ as

$$R(t_1, t_2) \equiv R(\tau) = \langle X(t)X(t+\tau)\rangle \tag{4.33}$$

The auto-correlation function $R(\tau)$ quantifies *statistical similarity* (see Fig. 4.8).

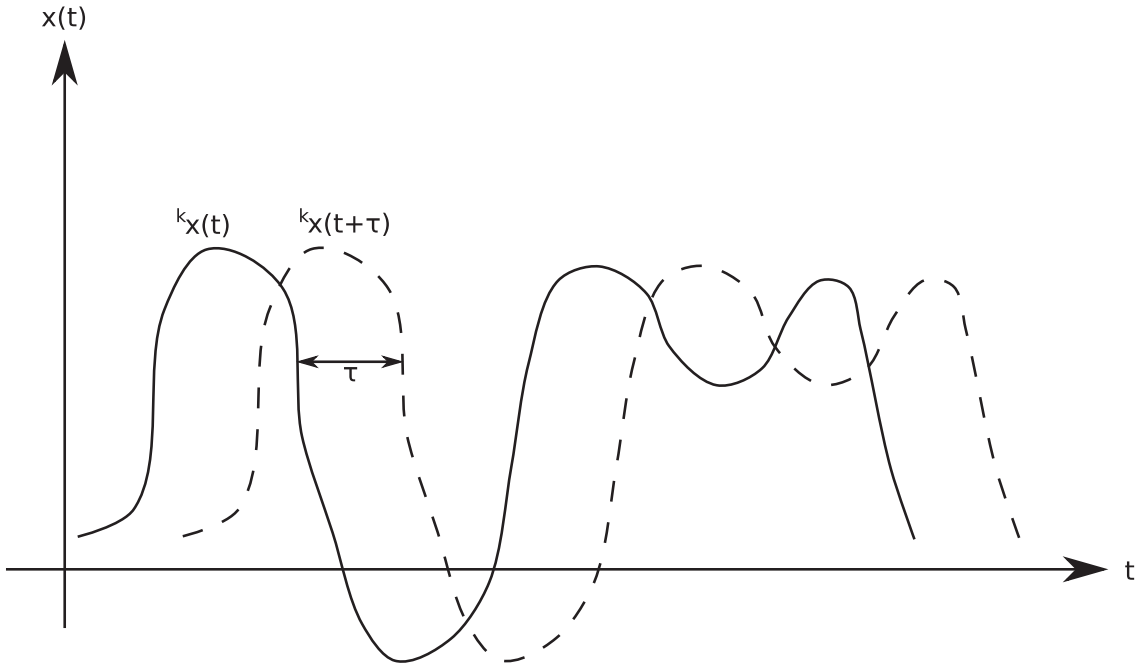

Figure 4.8: The autocorrelation function of a real random process $X(t)$ of zero mean serves as a measure of statistical similarity.